Structural Analysis of DARPin Mutants and Biochemical Characterization of ABC Transporters and SGT1

D i s s e r t a t i o n

zur Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät der

Universität Zürich

von

Michaela Andrea Kramer

aus

Deutschland

Promotionskomitee

Prof. Dr. Markus G. Grütter (Vorsitz & Leitung)

Prof. Dr. Raimund Dutzler

PD Dr. Ilian Jelezarov

Zürich, 2010

Table of content

Summary

My PhD thesis focuses on the structure analysis of proteins by X-ray crystallography, namely bacterial membrane proteins from the family of ABC transporters, the inflammasome adaptor protein SGT1 and Designed Ankyrin Repeat Proteins. These three topics will be summarized below.

Several ABC-transporters of *Escherichia coli* have been selected for recombinant expression and purification. By screening of several constructs and expression conditions the ABC-transporters YojI, MclB and HlyB were expressed. The purification protocol for the microcin exporter YojI was optimized and it was shown that YojI forms a stable dimer. This protein was used for functional assays. The obtained results form the basis for an increased understanding of the function of ABC-transporters.

The human protein SGT1 was expressed and characterized by biophysical methods. This protein is involved in the pathogen defence reactionin the mammalian immunesystem and its planhomologue is involved in the hypersensitivity response. Since the full length protein SGT1 did not yield crystals, we used limited proteolysis to obtain fragments that might be amendable for crystallization. Several fragments were obtained, cloned and overexpressed. For the central CS-domain an X-ray structures became available during my PhD thesis, for the N-terminal TPR domain structures of homologues were elucidated. The C-terminal SGS-domain revealed a limited degree of folding. The gained information is useful for further immuno-assays in order to determine interaction partners.

Selected Designed Ankyrin Repeat Proteins (DARPins) were structurally analysed. DARPins are extremely stable repeat proteins having a vast field of application as inhibitors and as crystallization aids. Previous studies have shown that the C-terminal capping repeat is limiting the stablility of these molecules. Therefore the group of Prof. A. Plückthun designed new C-terminal caps that showed improved thermal stability. To analyse the structural determinants for this improved thermal stability the DARPins NI_1C mutant 4, NI_3C mutant 5 and NI_3C mutant 6 were crystallized. The crystal structures revealed an improved packing between the C-terminal capping repeat and the last internal repeat explaining the improved thermal stabilityof the new design.

Zusammenfassung

Meine Dissertation beschäftigt sich mit der Röntgenstrukturanalyse von Proteinen. Zu diesen gehörten Mitglieder der Familie der ABC-transporter, das Inflammasome-Adapter-Protein SGT1 und den sogenannte Designed Ankyrin Repeat Proteins, welche folgend kurz beschrieben werden.

Eine Auswahl ABC-Transporter aus dem Organismus *Escherichia coli* wurde selektioniert, um eine einfache Expression zu gewährleisten. Bei dem Screening von verschiedenen Konstrukte und Expressionsbedingungen konnten die ABC-transporter YojI, MclB und HlyB expremiert werden. Es wurde gezeigt, dass der Microcin-exproter YojI ein stabiles Dimer bildet welcher für funktionelle Analysen genutzt wurde. Die hieraus hervorgegangenen Resultate werden eine Basis für das Verständnis der Funktion der ABC-Transporter von Nutzen sein.

Das humane Protein SGT1 wurde aufgereingt und durch biophysikalische Messmethoden charakterisiert. Dieses Protein ist in der Abwehrreaktion von Pathogenen in der Immunantwort des Säugetiers und sein pflanzliches Homolog in der so genannten Hypersensitivitätsreaktion beteiligt. Da das Protein SGT1 keine Kristalle bildete, wurden mit Hilfe von limitierten Proteolyseversuchen Fragmente bestimmt, welche die Kristallbildung vereinfachen sollen. Die dabei erhaltenen Fragmente wurden kloniert und aufgereinigt. Während dieser Arbeit wurde eine Struktur der mittleren CS-Untereinheit mittels Röntgenstrukturanalyse bestimmt, für die amino-terminale TPR-Untereinheit wurden Strukturen von Homologen bekannt. Die carboxy-terminale SGS-Unterinheit zeigte einen beschränkten Anteil an Faltung. Die erhaltenen Informationen können für die Ermittlung von Interaktionspartnern durch Immunoassays dienen.

Ausgewählte Designed Ankyrin Repeat Proteins (DARPins) wurden strukturbiologisch analysiert. DARPins sind extrem stabile Proteine welche aus sich wiederholenden Aminosäuresequenzen bestehen. Diese Proteine zeichnen sich durch eine weit gefächerte Anwendungsvielfalt aus, wie zum Beispiel ihr Nutzen als Inhibitoren und als Kristallisationshilfen. Vorangegangene Studien zeigten, dass die C-terminal Wiederholung die Stabilität dieser Moleküle limitiert. Deshalb entwarf die Gruppe von Prof. A. Plückthun diese C-terminale Wiederholung neu, ihre Resultate wiesen erhöhten thermischen Stabilitäten auf. Um die strukturellen Faktoren für diese verbesserte Stabilitäten zu bestimmen, wurden die DARPins NI_1C mutant 4, NI_3C mutant 5 and NI_3C mutant 6 kristallisiert. Die Röntgenkristallstrukturen zeigten eine dichtere Packung der carboxy-terminalen Wiederholung zu der letzten internen Wiederholung. Dieses Verhalten erklärt die verbesserte thermische Stabilität des neuen Designs.

List of abbreviations

α2β2 motif	motif comprising two antiparallel **α** -helices and a **β**-turn
ABC transporter	**A**TP **B**inding **C**assette transporters
AcrB	**Acr**iflavine resistance protein **B**
ADP	**A**denosine-**D**i-**P**hosphate
Amp	**Amp**icillin
AMP-PNP	**A**denylyl-imidodi**P**hosphate
APH	**A**minoglycoside **Ph**osphotransferase
ASC	**A**poptosis-associated **S**pe**c**k-like protein
ASU	**As**ymmetric **u**nit
At	*Arabidopsis thaliana*
ATP	**A**denosine-**T**ri-**P**hosphate
BIR	**B**aculoviral **I**nhibitor of apoptosis **R**epeats
CARD	**Ca**spase **R**ecruitment **D**omain
CC domain	**C**oiled-**C**oiled domain
C-cap	**C**arboxy-terminal **cap**ping repeat
CD	**C**ircular **D**ichroism
CFTR	**C**ystic **F**ibrosis **T**ransmembrane Conductance **R**egulator
CHAPS	3-[(3-**Ch**olamidopropyl)-dimethyl**A**mmino]-1 **P**ropanesulfonate
CHORD	**C**ysteine and **H**istine **R**ich **D**omain containing protein
CIITA	MHC **C**lass **II** **T**rans**A**ctivator
ColV	**Col**icin **V**
CS domain	**C**HORD and **S**GT1 motif containing domain
C-terminus	**C**arboxy-terminus
CV	**C**olumn **V**olume
Cymal 6	**6**-**Cy**clohexyl-1-hexyl-β-D-**Mal**toside
Cymal 7	**7**-**Cy**clohexyl-1-hexyl-β-D-**Mal**toside
DARPin	**D**esigned **A**nkyrin **R**epeat **P**rote**in**
DDM	n-**Dod**ecyl-β-**M**altopyranoside
DM	n-**D**ecyl-β-**M**altopyranoside
DNA	**D**eoxyribo**n**ucleic **a**cid
DTT	**Dit**hio**t**hreitol
E. coli	*Escherichia coli*
EDTA	**E**thylene **D**iamine Tetra **A**cidic acid
EPR spectroscopy	**E**lectron **Par**amagnetic spectroscopy
ESI	**E**lectro**S**pray **I**onization
Fos 12	n-**Dodecylphos**phocholine

Fos 14	n-**Tetradecylphos**phocholine
Fos 16	n-**Hexadecylphos**phocholine
GABPβ1	**G**uanine-**A**denine **B**inding **P**rotein**β1**
HEPES	4-(2-**H**ydroxy**e**thyl)-1-**P**iperazine**E**thane**S**ulfonic acid
HlyB	ABC transporter exporting **H**aemo**ly**sin A
HR	**H**ypersensitivity **R**esponse
Hs	***H**omo **s**apiens*
HSP90	**H**eat **S**hock **P**rotein **90**
Hv	***H**ordium **v**ulgare*
ICE	**I**nterleukin-1β **c**onverting **e**nzyme
IL	**I**nter**l**eukin
IM	**I**nner **M**embrane
IPTG	**Iso**propyl-β-D-**T**hio**G**alactopyranoside
IR	**I**nternal **R**epeat
Kan	**Kan**amycin
kDa	**k**ilo **Da**lton
LB media	**L**uria-**B**ertani media
LDAO	**L**auryl-**D**imethyl**A**mino-**O**xide
LPS	**L**ipo**P**oly**S**accharide
LRR	**L**eucine **R**ich **R**epeat
MALDI	**M**atrix **A**ssisted **L**aser **D**esorption **I**onizaton
MC RMSD	**M**ain **C**hain **R**oot **M**ean **S**quare **D**eviation
Mcc	**Mic**ro**c**in
MchF	**Mic**rocin exporting protein **F** exporting the microcin **H**45
McjD	**Mic**rocin exporting protein **D** exporting the microcin **D**
MclB	**Mic**rocin exporting protein **B** exporting the microcin **L**
MD	**M**olecular **D**ynamics
MDP	**M**uranyl**dip**eptide
MDR	**M**ulti **D**rug **R**esistance transporter
MFP	**M**embrane **F**usion **P**rotein
MHC	**M**ajor **H**istocompatibility **C**omplex
MS	**M**ass **S**pectroscopy
MW	**M**olecular **W**eight
NACHT domain	**NA**IP, **C**IITA, **H**ET-E and **T**P1 domain
NAIP	**N**LR family, **a**poptosis **i**nhibitory **p**rotein
NALP	**NA**CHT-, **L**RR- & **P**yrin-domain containing proteins
NBD	**N**ucleotide **B**inding **D**omains
NI1C	DARPin containing 1 internal repeat
NI3C	DARPin containing 3 internal repeats
Ni-IMAC column	**Ni**ckel-**I**mmobilized **M**etal **A**ffinity **C**hromatography

NLR	**N**BD **L**RR containing **R**eceptors
NLRA	**NLR** family containing **A**cidic Domains
NLRB	**NLR** family containing **B**IR Domains
NLRC	**NLR** family containing **C**ARD Domains
NLRP	**NLR** family containing **P**yrin Domains
NLRX	**NLR** family without strong homology to N-terminal domains of other NLRs
NMR spectroscopy	**N**uclear **M**agnet **R**esonance spectroscopy
NOD	**N**ucleotide-binding **O**ligomerization **D**omain containing protein
NOESY	**N**uclear **O**verhauser **E**nhancement **S**pectroscop**y**
N-terminal	amino-terminal
OD	**O**ptical **D**ensity
OG	n-**O**ctyl-β-**G**lucopyranoside
OM	**O**uter **M**embrane
OM	n-**O**ctyl-β-**M**altopyranoside
Os	*Oryza sativa*
OTG	n-**O**ctyl-β-D-**T**hio**G**lucopyranoside
PAMP	**P**athogen-**A**ssociated **M**olecular **P**atterns
PBP	**P**eriplasmic **B**inding **P**rotein
PDB	**P**rotein **D**ata **B**ank
PGN	Peptido**G**lyca**n**
Pgp	**P**-**g**lyco**p**rotein
Plk-1	**P**olo-**l**ike **k**inase **1**
PRR	**P**athogen-**R**ecognition **R**eceptor
PYD	**PY**rin **D**omain
RAR1	**R**equired for ML**A**12 **R**esistance protein
RIP	**R**eceptor **I**nteracting **P**rotein
RNA	**R**ibo**N**ucleic **A**cid
RP	**R**epeat **p**roteins
Sc	*Saccharomyces cerevisiae*
SC	**S**hape **C**omplementarity
scFabs	**s**ingle **c**hain **a**nti**b**ody **F**ragments
SD	**S**tandard **D**eviation
SDS	**S**odium **D**odecyl **S**ulphate
SDS-PAGE	SDS-**P**oly **A**crylamide **G**el **E**lectrophoresis
SEC	**S**ize **E**xclusion **C**hromatography
SGS domain	**SG**T1 **S**pecific domain
SGT1	suppressor of the G2 allele of skp1
SLS	**S**tatic **L**ight **S**cattering
TAP	**T**ransporter associated with **A**ntigen **P**resentation

TCEP	**T**ris(2-**C**arboxy**E**thyl)**P**hosphine
Tet	**Tet**racyclin
TLR	**T**oll-**L**ike **R**eceptors
TMD	**T**rans**M**embrane **D**omain
TPR	**T**etratrico **P**eptide **R**epeat
Tris	**Tris**(hydroxymethyl)-aminomethan
UDM	n-**U**n**D**ecyl-β-**M**altopyranoside
VR	**V**ariable **R**egion

Chapter A - ABC transporter

A.I Introduction

A.I.1 General Overview

The transport of solutes across the biological membrane is an essential prerequisite for the survival and functioning of both prokaryotic and eukaryotic cells. Approximately a third of the proteins of a cell are membrane spanning proteins, of which one important family are the ATP binding cassette (ABC) transporters. They transport solutes over the cellular membrane under the consumption of ATP.

Members of this family can function as channels, importers or exporters. Up to now ABC importers are only known in prokaryotes, whereas ABC exporters are characterized in both, prokaryotes and eukaryotes (Davidson 2008).

Their activities are required for various processes essential to cells, such as the uptake of nutrients (examples are sugars, ions, amino acids), osmosensing and -regulation as well as antigen processing. Multidrug resistance, a major clinical problem in the treatment of human cancers, as well as bacterial and fungal infections, are mediated by this transporter class. In addition to this, the dysfunction of an ABC transporter forms the molecular basis of several inherited human diseases, including cystic fibrosis, bile salt disorders and adrenoleukodystrophy (Dean and Allikmets 1995). This protein class is therefore of high biomedical interest.

So far about 48 human ABC transporters have been characterized (Dean, Rzhetsky et al. 2001). They are classified in 6 different subfamilies (ABCA, ABCB, ABCC, ABCD, ABCE/ABCF and ABCG). The subfamily ABCA can be subdivided in two phylogenetic groups. The first group comprises ABCA1, ABCA2, ABCA3, ABCA4, ABCA7, ABCA12, ABCA13; the second group contains ABCA5, ABCA6, ABCA 8, ABCA9 and ABCA10, all clustered on the same chromosome.

A well known ABC transporter is the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR), a chloride ion channel, belonging to the subfamily ABCC (Quinton 1999). Members of the subfamily ABCB are the well studied multidrug resistance transporter ABCB1 also named P-glycoprotein or Pgp (Rosenberg, Callaghan et al. 1997) and TAP involved in antigen presentation (Muller, Ebensperger et al. 1994). ABCA1 is known for cholesterol transport disorders and high-density lipoprotein biosynthesis (Dean, Rzhetsky et al. 2001).

A.I.2 Domain Arrangement in the different ABC-transporters

ABC transporters are composed of two transmembrane domains (TMD) being responsible for the substrate transport across the membrane and two nucleotide binding domains (NBD), that are located in the cell interior and fuel by binding and hydrolyzing ATP (Figure A.1).

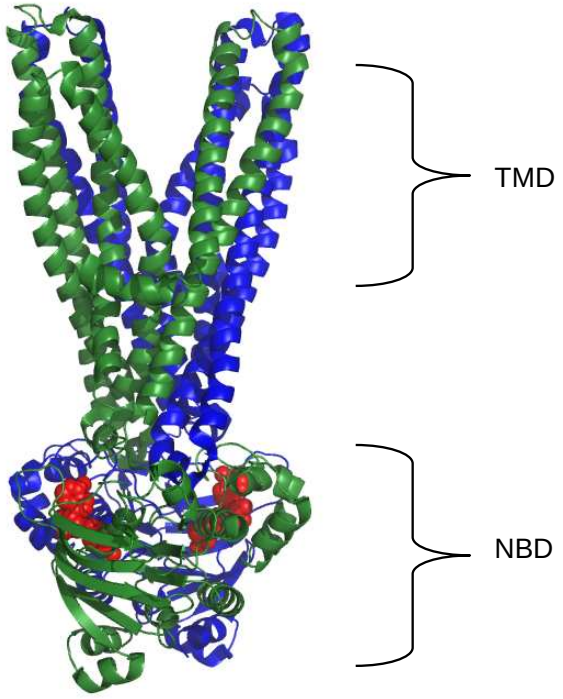


Figure A.1) The structure of the ABC transporter Sav1866 of *Staphylococcus aureus* in its outward-facing conformation representing the ATP-bound state (Dawson and Locher 2006). The two monomers are coloured in green and blue respectively. The bound nucleotide ADP is shown in red spheres.

Figure A.2 displays different domain organizations observed in the ABC transporter superfamily (reviewed by (Locher 2009), (Higgins 1992), (Rees, Johnson et al. 2009)).

In importers, the four domains assemble as four independent polypeptide chains to form the functional transporter (Figure A.2a). In addition, a periplasmic binding protein (PBP) is required that captures the substrate in the periplasm and delivers it to the ABC importer. Examples for ABC importers are the vitamine B12 uptake system BtuCD, which is the first solved structure of an ABC transporter (Locher, Lee et al. 2002), or the maltose uptake system of *E.coli* (Oldham, Khare et al. 2007). The TMDs of the structurally and functionally diverse ABC importers consist of 5 to 10 transmembrane helices (Rees, Johnson et al. 2009). The TMD of exporters in contrast always consist of 6 transmembrane helices and are fused to a NBD forming a half transporter (Figure A.2c). Two half transporters

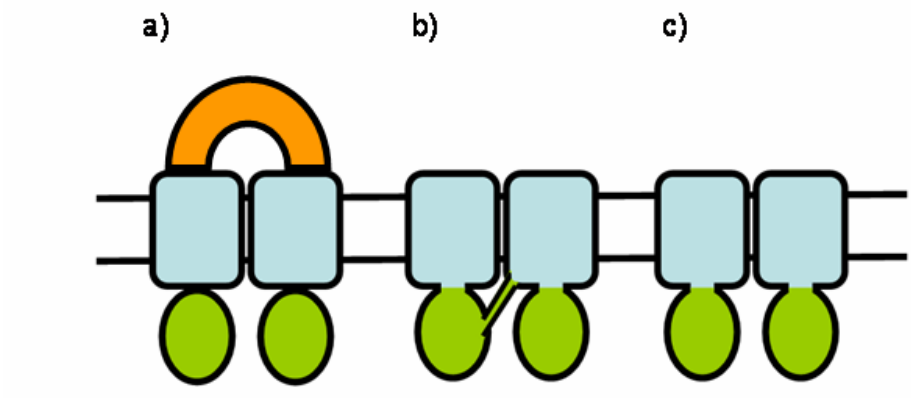


Figure A.2) Different structural domain organization found in ABC-transporters. The cytosolic NBD is coloured in green, the TMD is coloured in light blue. The PBP is coloured in orange. **a)** All domains are expressed as single polypeptides, as found in ABC-importers having a PBP; **b)** All domains are fused to one polypeptide chain; **c)** two half transporter consisting of a TMD-NBD fusion form the functional unit.

homo- or heterodimerize to form the active ABC transporter, like in the case of the recently solved X-ray structures of Sav1866 of *Staphylococcus aureus* (Dawson and Locher 2006) and the lipid flippase MsbA (Ward, Reyes et al. 2007). Also the transporters discussed in this work, the bacterial ABC transporter YojI, and the microcin exporters MclB and MchF share this architecture.

In eukaryotes, all four domains are usually fused to a single polypeptide chain, as for instance in the prominent examples of the human multidrug resistance transporter MDR and CFTR (Higgins 1992) (Figure A.2b) .

ABC dimers devoid of TMDs also exist and are involved in DNA repair (e.g., RecA (Story and Steitz 1992) and Rad50 (Hopfner, Karcher et al. 2000)).

A.I.3 Structural Motifs in ABC transporters

The NBDs display a high degree of homology (up to 30 % sequence identity (Holland and Blight 1999)) and contain a number of conserved sequence motives that are required for the binding and hydrolysis of ATP (Figure A.3).

The arrangement of the two sub domains of the NBD are assembled as two lobes standing rectangular on each other. The first domain is the so called RecA like domain (Walker, Saraste et al. 1982). The second one is named helical sub-domain and consists of three to four helices. In the presence of ATP, the NBDs dimerize in a head-to-tail conformation; the RecA-like domain of one NBD is facing the helical sub-domain of the counterpart and thereby forms a catalytic site involved in ATP hydrolysis.

The helical subdomain and the RecA-like domain are linked by the so called Q-loop, containing a highly conserved glutamine: This loop interacts with the TMD of the full transporter and assures the transmission of the energy obtained by ATP hydrolysis to a structural reorganization of the transmembrane domain, resulting in a different conformation in the translocation path of the transported substrate.

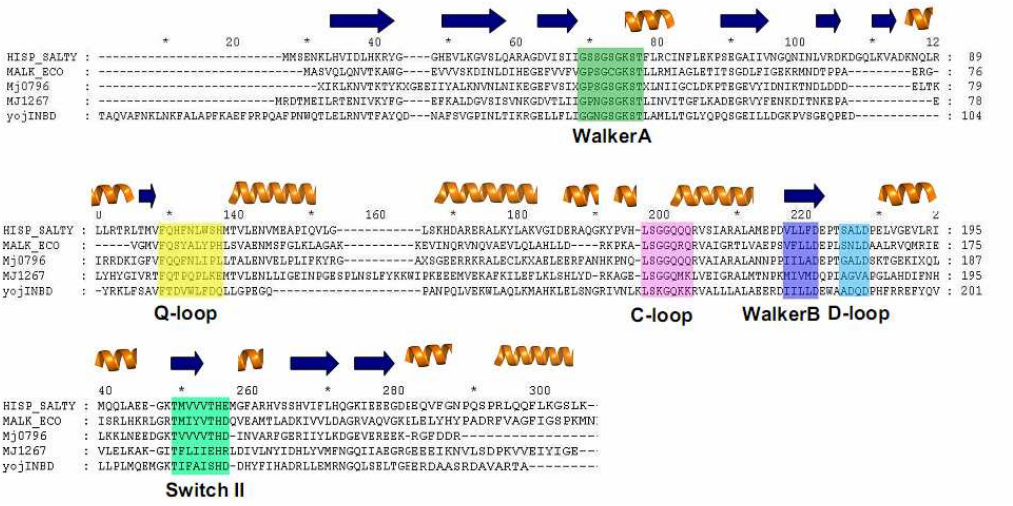


Figure A.3) Structural alignment of the NBD of HisP of *S. typhimurium*, MalK from *E. coli*, MJ0796 and MJ1267 from *Methanococcus jannaschii* and YojI of *E. coli*. The characteristic motifs are coloured. Predicted secondary structure elements, as predicted by JPRED for the NBD of YojI are depicted.

Before the hydrolysis reaction takes place, ATP is bound in a sandwich-like fashion between the Walker A motif (GxxGxGKS/T, x is any amino acid), of the cis-monomer and the ABC signature motif, the so called LSGGQ motif, of the trans-monomer. The cofactor Mg^{2+} is coordinated by the aspartate of the Walker B motif and is crucial for the hydrolysis reaction. The lysine of the Walker A motif, the histidine of the H-loop and the glutamate of the Walker B motif form a charge network with the phosphate moiety of ATP and play a key role in ATP hydrolysis.

A.I.4 The ATPase cycle

Despite extensive studies on isolated NBDs, and the crystal structures of several full transporters, the catalytic cycle of ATP hydrolysis and the energy coupling to the substrate transport in the TMDs are still controversially discussed in the literature.

Two models are debated how the ATP hydrolysis is coupled to the transport reaction. The first model was proposed by Higgins and co-workers (Higgins and Linton 2004) in which the ATP binding and hydrolysis induces the formation and dissociation of

the NBD dimer, respectively. The NBD dimerization is proposed as the regulatory switch, being responsible for conformational changes in the TMDs that consequently decrease the affinity for the binding of the substrate and allows for drug release to the cell exterior.

A second model is based on observations with vanadate induced trapping of ATP in P-glycoprotein. These experiments gave evidence that only 1 ATP can be trapped in one Pgp dimer by vanadate (called ATP occlusion) and suggested that only 1 ATP can be hydrolyzed at the same time (Urbatsch, Sankaran et al. 1995). The occlusion of the ATP was shown to induce the conformational switch, resulting in the high to low affinity transition of the TMDs to the substrate, which is consequently transported out of the cell (Sauna and Ambudkar 2007), as depicted in Figure A.4 (reviewed by (Seeger and van Veen 2009)).

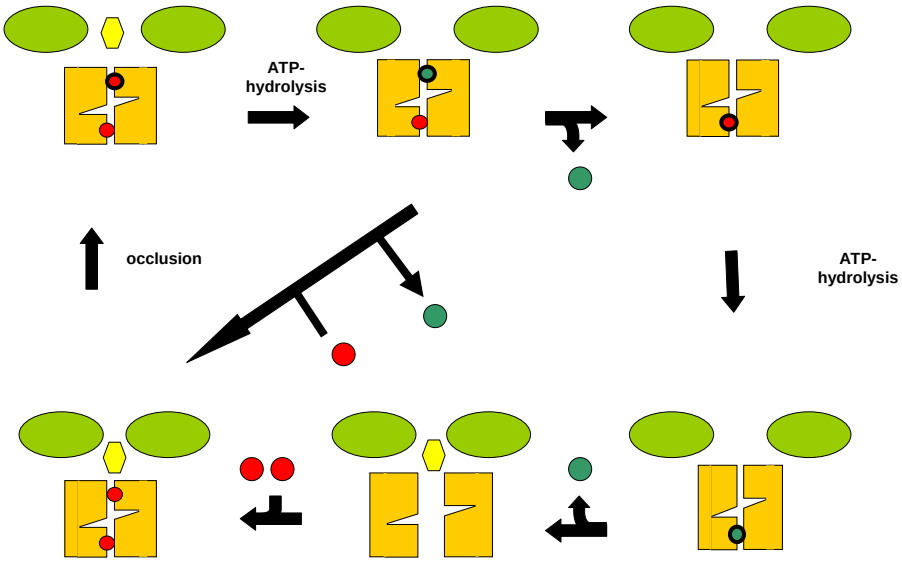


Figure A.4) The occlusion induced switch model as established by senior and co workers. The occlusion of ATP is necessary for the hydrolysis reaction and decreases the affinity for the substrate, which is then liberated. Subsequently ADP is released. Either the second ATP is hydrolysed or a new ATP molecule can be bound. Substrate can bind again, so that a new transport reaction can be started (red = ATP, green=ADP).

Crystal structures revealed the conformational changes taking place during the transport cycle of ABC exporters.

In the crystal structure of Sav1866, crystallized in an outward facing conformation in the presence of ADP or AMP-PNP (Dawson and Locher 2006) the membrane helices are split up into two wings. The substrate is proposed to escape either to the periplasm or to the outer leaflet of the membrane, whereas the NBDs are displaying tight dimer interactions. Two coupling helices form the transmission interface between the TMDs to the NBDs and assure the transmission of the structural rearrangement from the NBDs to the TMDs by non-covalent interactions (Dawson and Locher 2006).

More information on the possible conformational repertoire of ABC exporters is available for the lipid flippase MsbA that was crystallized in three different conformations: an opened apo form, a closed apo-form and a closed nucleotide-bound form (Ward, Reyes et al. 2007). The latter one was crystallized either in the presence of Mg-V_i-ADP or with two AMP-PNP (Borbat, Surendhran et al. 2007). In the absence of nucleotides, the transporter adopts an inward-facing conformation and the substrate is proposed to be accommodated either from the cytoplasm or the inner leaflet of the cell membrane. In the closed apo-form, the NBDs are in closer proximity whereas in the closed nucleotide bound conformation, the NBDs dimerize and the TMDs are in an outward facing conformation, comparable to the structure of Sav1866. EPR spectroscopy and site-directed spin-labeling are in support of the conformations of MsbA found in the X-ray structures (Borbat, Surendhran et al. 2007).

These conformational changes are in agreement with the alternate access and release model that was also postulated for Sav1866 (Dawson and Locher 2006).

A.I.5 Microcin-exporter

Microcins as well as colicins are bacteriocins produced by *E. coli*. In contrast to colicins, microcins are non-lethal to their producing strain but they inhibit the

growth of closely related strains. These microcins have a molecular weight of less than 10 kDa (reviewed by (Pons, Delalande et al. 2004)) and can be divided in two classes, according to their size and modifications.

Class I microcins (Mcc) are post-translationally modified and are targeted on intracellular processes. They have a molecular mass of less than 5 kDa. Mcc B17, Mcc C7, Mcc J25 and possibly Mcc D93 all belong to this group.

To the second class belong microcins with a molecular mass between 7-10 kDa. Mcc E492, Mcc L, Mcc H47 and Mcc V are firmly apart of the class II group. (reviewed by (Pons, Lanneluc et al. 2002; Pons, Delalande et al. 2004).

In general, four clustered genes, grouped in one or two operons encode the machinery for the production and export of microcins. The genes encode the precursor peptide, an immunity protein conferring resistance against the microcine, the microcin exporter and the membrane fusion protein (MFP). It is believed that a tripartite assembly assures the export of the substrate out of the bacterial cell, as seen for the ColV secretory complex (Figure A.5) ((Azpiroz, Rodriguez et al. 2001; Pons, Delalande et al. 2004). The transporter protein itself is located in the inner membrane (IM) of *E. coli*. For the secretion of a toxin, it requires TolC, an outer membrane protein, mediating the transport of the toxin through the outer membrane (OM) and the MFP creating an IM-OM bridge (Hwang, Zhong et al. 1997).

The microcin Mcc J25 belongs to the small class I microcins. It is synthesized from a 58 amino acid long precursor and after processing, the active toxin is 21 amino acids long and forms a so called lariat ring as seen in Figure A.5b (Blond, Cheminant et al. 2001). The microcin J25 is plasmid-encoded together with its associated ABC transporter McjD (Delgado, Vincent et al. 2005). However, Mcc J25 has been shown to be transported by YojI as well (Delgado, Vincent et al. 2005) .

The microcins MccH47 and MccL are two examples for class II microcins. They possess a leader sequence, which is cleaved upon externalization. The dedicated microcin exporters MclB and MchF have a protease domain which is fused to

a) **b)**

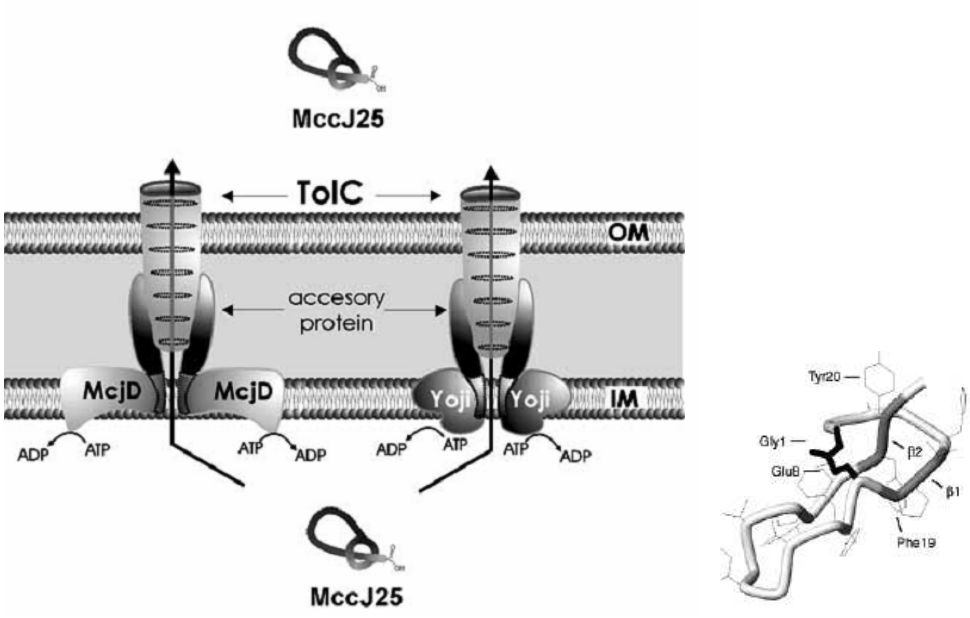


Figure A.5) a) The export of the microcinJ25 (Vincent and Morero 2009). The MccJ25 Synthase processes the MccJ25 precursor, in order to obtain its characteristic lasso structure. This microcin can be exported by the ABC-transporter McjD or YojI. The accessory proteins are unknown. **b)** The microcine J25 as determined by NMR (Blond, Cheminant et al. 2001)

the N-terminus of the transporter, assuring the leader sequence cleavage (Havarstein, Diep et al. 1995).

The ABC transporter focused on in this PhD thesis is YojI from *E. coli* (Kobayashi, Nishino et al. 2001). It is a typical homodimeric ABC exporter mediating the export of the microcin J25. So far, the MFP interacting with YojI is still unknown. Additional ABC-transporters were examed in this thesis, namely the exporter of Haemolysin A (HlyB), and the bacteriocin exporters of the class II microcins L (MclB) and H47 (MchF) from *E. coli*.

A.II Results

A.II.1 Selection of the transporters

The main goal of this project of the PhD thesis was to purify ABC transporters at high yield that are stable in detergent solution in order to determine the three dimensional structure by X-ray crystallography. Since the purification and crystallization is known to be tedious and sometimes even impossible for a multitude of membrane proteins, several ABC transporters were examined in this study.

The focus was on ABC exporters, in particular the hemolysin exporter HlyB and the microcin exporters MclB, MchF and YojI. The aim was to choose half transporters, having the NBD fused to the C-terminus of the transmembrane domain in order to facilitate a straightforward cloning and purification procedure. The transporter MclB and MchF have an additional N-terminal proteolytic domain fused to the N-terminus, as depicted in Figure A.6. This domain cleaves the double-glycine leader of class II microcines.

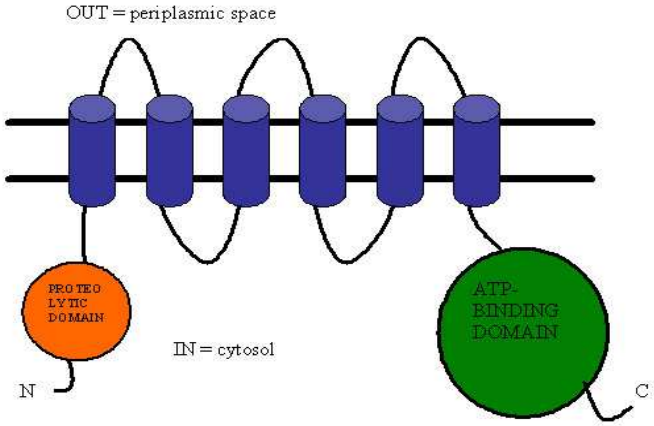


Figure A.6) Scheme of the ABC half transporter MclB. It possesses 6 transmembrane helices. The NBD, is fused to the C-terminus. The proteolytic domain is located at the N-terminus.

The genes were cloned by standard techniques in a subset of vectors. Cloning of *mchE/F* was unsuccessful presumable due to toxicity problems caused by the expression of this protein in *E. coli*.

The constructs used in this chapter are listed in Table A.1, showing also whether the respective transporter was expressed and purified.

Table A.1) Expression constructs of ABC-transporters used in this project. The expression conditions and the expression level of the respective protein are enlisted.

Gene (*E. coli* strain of origin)	**Vector**	**Remark**	**Expression hosts**	**Protein production (Western blot or dot blot)**	**Protein purification**
mclB (LR05)	pBAD24	C-terminal His6-tag	*E. coli* BL21(DE3) BL21(DE3)-AI C41(DE3) C43(DE3)	-	Not purified
mclB (LR05)	pET28	C-terminal His6-tag	*E. coli* BL21(DE3) C41(DE3) C43(DE3) DH5α	+ Dot Blot weak band on Western Blot	Not purified
mclB (LR05)	pET28	N-terminal His6-tag	*E. coli* BL21(DE3) C41(DE3) C43(DE3)	-	Not purified
mclB (LR05)	pET21	C-terminal His6-tag	*E. coli* BL21(DE3) BL21(DE3)plys C41(DE3) C43(DE3) DH5α	-	Not purified
mclB (LR05)	pET21	N-terminal His6-tag	*E. coli* BL21(DE3) C41(DE3) C43(DE3)	-	Not purified
hlyB (CFT07)	pET21	C-terminal His6-tag	*E. coli* BL21(DE3) C41(DE3) C43(DE3)	+ Western Blot	Low yield, after Ni-NTA

Gene (*E. coli* strain of origin)	Vector	Remark	Expression hosts	Protein production (Western blot or dot blot)	Yield of purified protein
hlyB (536)	pET21	C-terminal His6-tag	*E. coli* BL21(DE3) C41(DE3) C43(DE3)	-	Not Purified
hlyB (536)	pET21	N-terminal His6-tag	*E. coli* BL21(DE3) C41(DE3) C43(DE3)	-	Purified
hlyB (536)	pET28	C-terminal His6-tag	*E. coli* BL21(DE3) C41(DE3) C43(DE3)	+ Dot Blot + Western Blot	Low yield, after Ni-IMAC
hlyB (536)	pET28	N-terminal His6-tag	*E. coli* BL21(DE3) C41(DE3) C43(DE3)	-	Not purified
hlyB (CFT07)	pBAD24	C-terminal His6-tag	*E. coli* BL21(DE3)	-	Not purified
hlyB (536)	pBAD24	C-terminal His6-tag	*E. coli* BL21(DE3)	-	Not purified
mchE/F (CA58)	pBAD24	C-terminal His6-tag	*E. coli* BL21(DE3)-AI	-	Not purified
mchE/F (CA58)	pET21	C-terminal His6-tag	*E. coli* BL21(DE3)	-	Not purified
YojI (CFT07)	pET28	C-terminal His6-tag	*E. coli* BL21(DE3)	+ Dot Blot + Western Blot	Purified

For the three ABC transporters YojI, MclB and HlyB, protein expression could be demonstrated by Dot blotting and Western blotting (Figures A.7 and A.8). For YojI, protein production could be observed under all tested conditions, whereas only a weak signal was detectable for MclB after 4h expression time at a temperature of 30°C (Figure A.7a). The expression levels of YojI were sufficient to purify the protein to homogeneity (see below). For MclB, a band at an apparent molecular weight of 70 kDa was detected which agrees with its expecected molecular weight of 76.8 kDa

(Figure A.7b). Likewise, protein expression could be demonstrated for HlyB with a number of constructs (Figure A.8). The expression yield of MclB however, was too low to allow for protein purification.

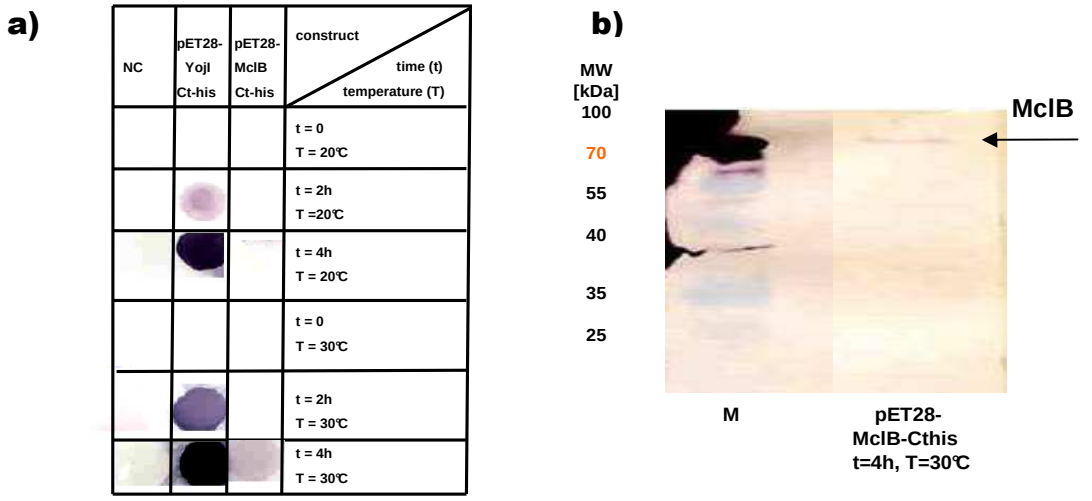
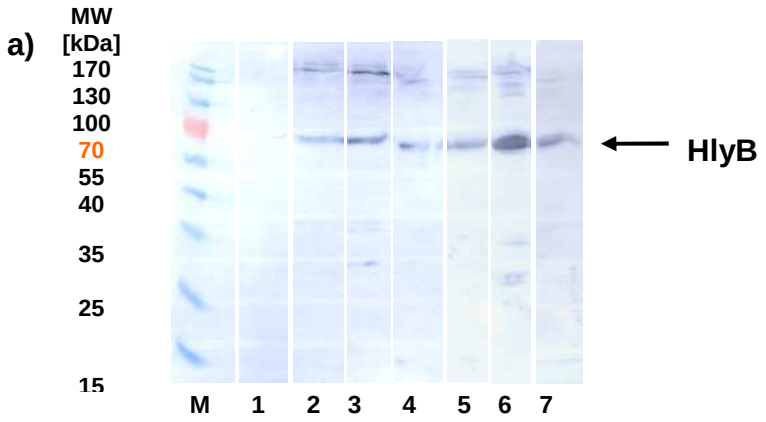


Figure A.7) Screening for expression conditions for the ABC-transporters MclB and YojI a) Dot Blot of culture samples expressing YojI or MclB as compared to a negative control (NC), which were taken after 0, 2 and 4h after induction at the temperatures of 20°C and 30°C. b) Western Blot analysis of cells expressing MclB taken at 4 h after induction at 30°C. The proteins were detected with a scFv antibody selective for the His-tag.



lane		lane			lane		
M	MW marker	2	pET28HlyB(CFT037)	t = 2h	5	pET28HlyB(CFT037)	t = 4h
1	NC t=0h	3	pET21HlyB(536)	t = 2h	6	pET21HlyB(536)	t = 4h
		4	pET28HlyB(536)	t = 2h	7	pET28HlyB(536)	t = 4h

Figure A.8) Western Blot analysis of HlyB expressed from different vectors in *E. coli* BL21(DE3).

A.II.2 Purification of the constructs of the ABC-transporter

YojI turned out to be the only expressing transporter that could be purified yielding reasonable amount of pure protein. Therefore it was decided to carry out a thorough screening on buffer compositions and detergents with the ultimate aim to obtain a high yield of stable, non-aggregated and functional dimeric protein.

A.II.2.a Optimization of the buffer conditions for YojI of *E.coli*

The initial purification protocol, using the buffer containing 20 mM NaCl pH 7.5, 0.03 % DDM yielded highly pure protein (Figure 9b). However, when the protein was separated by size exclusion chromatography (SEC), aggregation was observed (Figure A.9a). Notably, a peak corresponding to the Yoji dimer could be collected (Sennhauser 2004), that was prone to aggregation upon prolonged storage as judged by size exclusion chromatography (not shown).

Therefore, different buffers covering a pH range between pH 6.5 and pH 7.5 were used to purify the protein, in order to determine conditions under which the protein is less prone to aggregation (Table A.2).

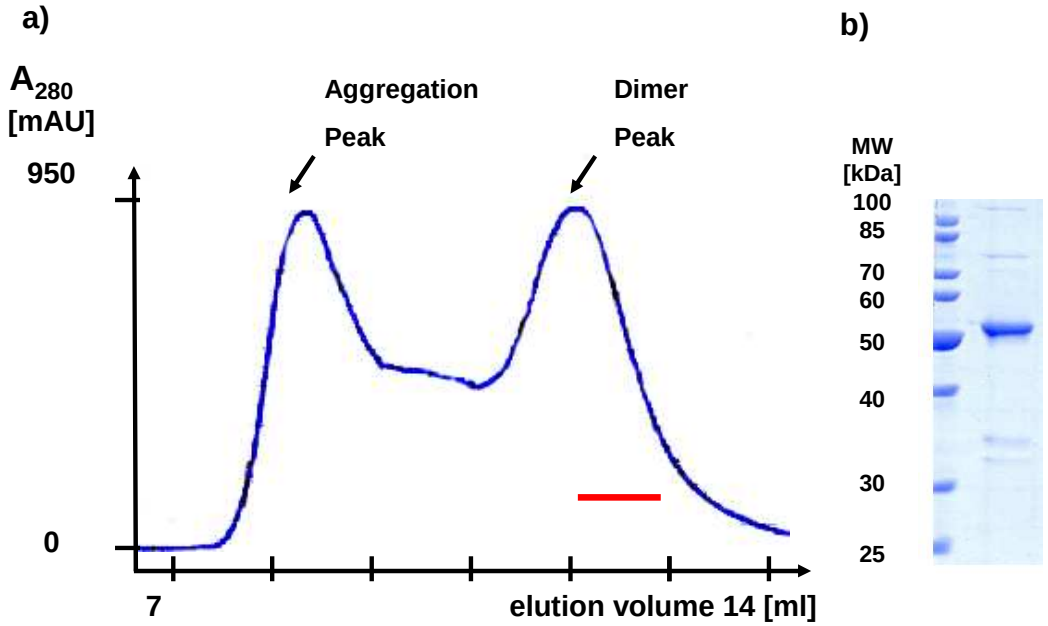


Figure A.9) Initial purification attempt of YojI in 20 mM Tris, 150 mM NaCl pH 7.5, 0.03 % DDM. **a)** Size exclusion chromatogram. **b)** SDS-PAGE analysis of the YojI fractions marked in red in pannel a)

Table A.2) Buffer screen using size exclusion chromatography in order to reduce the aggregation of YojI.

Buffer, pH, detergent (3X CMC)	**Ratio between dimer peak and aggregation peak (peak height) on SEC**
20 mM Tris, 7.5, DDM	40:60
20 mM Citrate, 6.4, DDM	80:20
20 mM HEPES 7.5, DDM	0:100
20 mM KPi, 7.5, DDM	0:100
20 mM KPi, 7.2, DDM	0:100
20 mM KPi, 7.0, DDM	20:80
20 mM KPi, 6.5, DDM	0:100
20 mM Citrate, 6.8, 500 mM NaCl, DDM	40:60
20 mM Citrate, 6.8, Glycerole Cymal 7	60:40
20 mM Citrate, 6.4 Cymal 7	80:20
20 mM Citrate, 7 Cymal 7	80:20

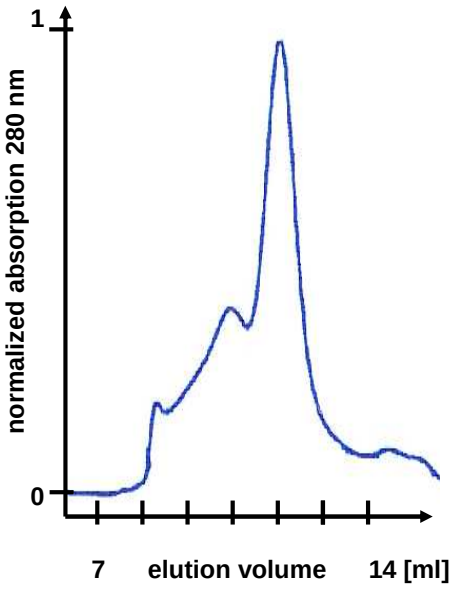


Figure A.10) Size exclusion chromatography of the ABC-transporter YojI of *E. coli* using 20 mM Na-Citrate, 150 mM NaCl, 0.03 % DDM pH 6.4

The purification yielding the best peak separation and the smallest amount of aggregated protein was obtained in 20 mM Citrate and 150 mM NaCl pH 6.4 (Figure A.10).

A.II.2.b Screening for detergents

Based on the results from the buffer screening, it was decided to work with the buffer yielding the best peak separation and the smallest aggregation peak, namely 20 mM Citrate pH 6.4, 150 mM NaCl supplemented with 0.03 % DDM.

In a next step, it was decided to screen for detergents suitable for the purification of YoijI. In a first approach different detergents were screened by solubilizing the protein YojI in DDM, a detergent known to be suitable for purification (Sennhauser 2004). After protein binding onto the Ni-IMAC column, DDM was exchanged for eight different detergents during the subsequent washing steps (Table A.3). Further analysis revealed that the detergents DDM, Cymal-7, Fos-12 and UDM are suitable for the purification of YojI.

Table A.3) Detergent screening for YojI. The protein was first solubilized in DDM which was exchanged for the tested detergent on the Ni-IMAC column.

Detergent	Ratio between dimer peak and aggregation peak (peak height) on SEC	Remarks
Cymal-6 0.028 %	70:30	
Cymal-7, 0.01 %	70:30	Peak between aggregate and dimer
Fos-16 0.00159 %		Precipitation
Fos-14 0.014 %	80:20	Oligomerization peak
Fos-12 0.15 %	60:40	Good peak separation
LDAO 0.069 %	60:40	Oligomerization peak
Thesit 2.7%	55:45	Poor peak separation
UDM, 0.05 %	20:80	Protein prone to aggregation when concentrated

With this subset of detergents, the purification was performed by using the respective detergent already in the solubilization step (Table A.4).

Table A.4) Purification in different detergents. The protein was solubilized and purified in the same detergent.

Detergent	Ratio of Dimer peak and aggregation peak (peak height) in SEC	Yield	Remarks
DDM, 0.03 %	60:40	1 mg / litre of culture	
Cymal-7, 0.01 %	70:30	0.5 mg / litre of culture	Additional peak between aggregate and dimer
Fos-12 0.15 %	60:40	1 mg / litre of culture	
UDM, 0.05 %	80:20	0.25-1 mg / litre of culture	Protein prone to aggregation when concentrated

The resulting size exclusion chromatogram and SDS-PAGE analysis are depicted in Figures A.13 and A.9.

Previous studies showed that the protein YojI did not show a high stability in the detergent DDM (data not shown). However, the isolated dimer peak of Yoji exhibited a high degree in stability for the selected detergents UDM and Cymal 7, as seen by reinjecting the peak after a period of one week (Figure A.14).

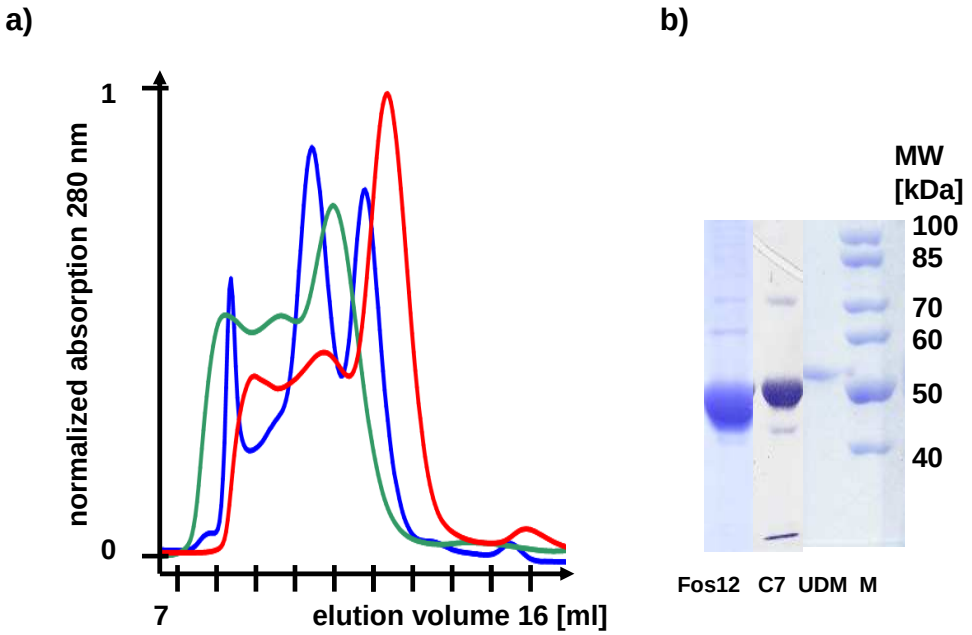


Figure A.13) Purification and SEC analysis of YojI in the detergents UDM, Cymal7 and Fos-12. **a**) SEC chromatogram of the last purification step of YojI in UDM (red), Cymal 7 (green) and Fos-12 (blue). **b)** SDS-PAGE analysis of the fractions of the last peak of the different detergents (M=MW marker, Fos12= Foscholine 12, C7=Cymal7, UDM=UDM).

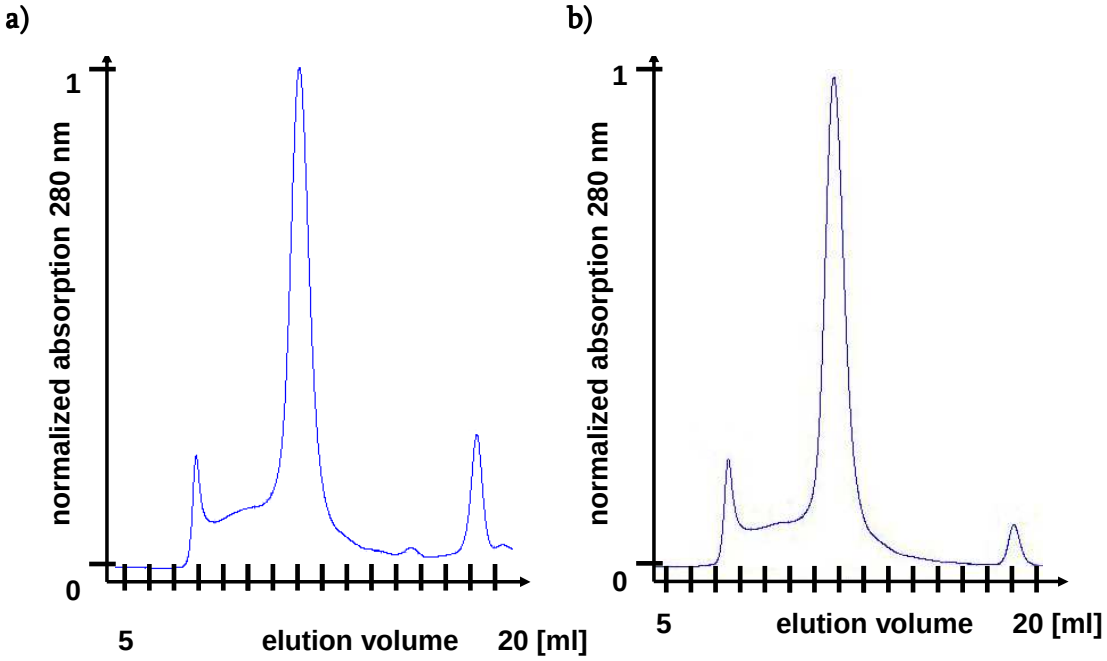


Figure A.14) SEC of the purified protein YojI after 7 days incubation at 4°C in the different detergents. **a)** UDM, **b)** Cymal 7.

Unfortunately, no basal ATP hydrolysis activity could be measured with purified YojI. Moreover, no binding of the transporter to ATP-linked agarose resin could be detected. Nevertheless, it was decided to proceed with crystallization experiments.

A.II.2.c Purification experiments of the ABC-transporter HlyB of E. coli

The ABC-transporter HlyB of the *E.coli* strain CFT037, expressed in the *E.coli* strain BL21(DE3) was purified according to the established protocol (Sennhauser 2004). However, only a small amount of protein could be purified, because the majority of the protein aggregated on the Ni-IMAC resin during the washing steps (Figure A.15). Trials to improve the yield using different detergents (DDM, DM and LDAO) or purifying HlyB originating from the *E.coli* strain 536, expressed in the *E. coli* strain BL21(DE3) were unsuccessful.

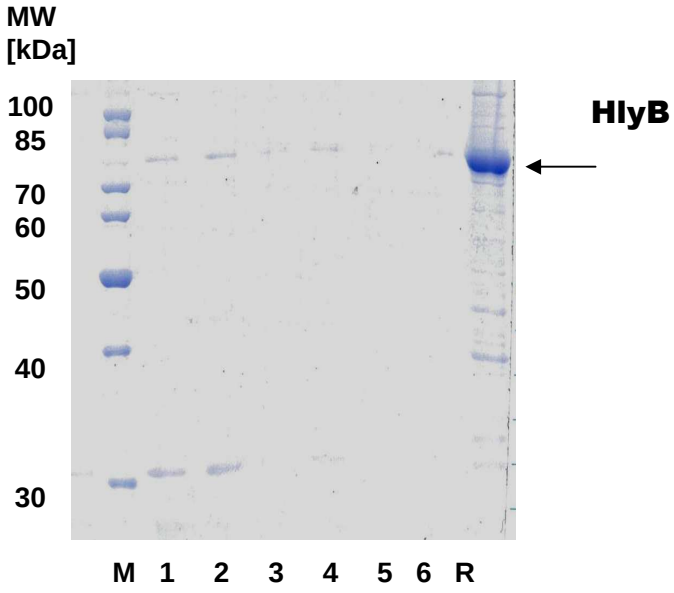


Figure A.15) SDS-PAGE analyis of a purification attempt of HlyB of the *E.coli* strain 536 expressed in the *E. coli* strain BL21(DE3) in 20 mM Tris, 150 mM NaCl pH 7.5, 0.03% DDM M= MW marker, Samples 1-6: Stepwise elution (1ml fractions) from the IMAC column with elution buffer containing 250 mM imidazole, R= loaded IMAC resin.

A.II.3. Crystallization experiments of the ABC-transporter YojI

Crystallization experiments of the protein YojI were done at various protein concentrations and buffer conditions, as listed in table 5. Prior of setting up the trials, the protein was fixed in different conformation, suitable for crystallization. The addition of ADP and vanadate, the ATP–analogue AMP-PNP and the cofactor $MgCl_2$ are known to arrest the NBD of the transporter in the different conformation states of the hydrolysis cycle. However, no crystals could be obtained.

Table A.5) List of the crystallization trials of the protein YojI. The detergent, concentration, additions prior to crystal set ups, as well as the screens are indicated.

Detergent	Concentrations	Additions prior to crystal set ups	Screens
Cymal 7	5-20 mg/ml	$MgCl_2$	McKinnon Screens 1-6
Cymal 7	5-20 mg/ml	Na_3VO_4, ADP, $MgCl_2$	McKinnon Screens 1-6
Cymal 7	5-20 mg/ml	ATP, $MgCl_2$	McKinnon Screens 1-6
DDM	10 mg/ml	$MgCl_2$	McKinnon Screens 1-6
DDM	10 mg/ml	Na_3VO_4, ADP, $MgCl_2$	McKinnon Screens 1-6
Fos-12	10 mg/ml		McKinnon Screens 1-6
Fos-12	10 mg/ml	$MgCl_2$	McKinnon Screens 1-6
Fos-12	10 mg/ml	$MgCl_2$, ATP	McKinnon Screens 1-6
Fos-12	10 mg/ml	Na_3VO_4, ADP, $MgCl_2$	McKinnon Screens 1-6
Fos-12	10 mg/ml	ADP	McKinnon Screens 1-6
UDM	10 mg/ml	$MgCl_2$	McKinnon Screens 1-6
UDM	10 mg/ml	$MgCl_2$, ATP	McKinnon Screens 1-6

Detergent	Concentrations	Additions prior to crystal set ups	Screens
UDM	10 mg/ml	Na_3VO_4, ADP, MgCl2	McKinnon Screens 1-6
DM	10 mg/ml	$MgCl_2$	McKinnon Screens 1-6
DM	10 mg/ml	$MgCl_2$, ATP	McKinnon Screens 1-6
DM	10 mg/ml	Na_3VO_4, ADP, $MgCl_2$	McKinnon Screens 1-6

A.II.4. Nucleotide Binding domains of the ABC-transporters YojI and HlyB

The coding sequences of the NBD of the protein YojI and HlyB were successfully cloned and over-expressed using the vector pET21 with a C-terminal His-tag, as seen in Table A.6.

Table A.6) Constructs used for characterization of the NBDs of ABC-transporters. The expression host, yield and activity are shown.

Gene	Vector	Remark	Expression host	Expression condition	Yield of purified protein	Acitivty
HlyB-NBD (536)	pET21	C-terminal His6-tag	*E. coli* BL21(DE3)	Grown in LB, 37°C 4h expression time	3 mg/ 200ml bacterial culture	Not active
YojI-NBD (CFT037)	pET21	C-terminal His6-tag	*E. coli* BL21(DE3)	Grown in LB, 37°C 4h expression time	3 mg/ 200ml bacterial culture	Not active, binding to ATP-agarose

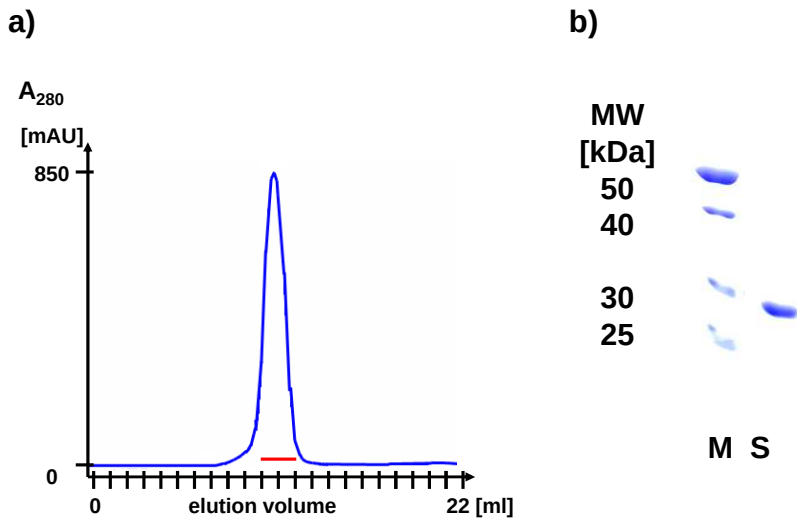


Figure A.16) Purification of the NBD of the transporter YojI. **a)** Size exclusion chromatogram depicts a monodisperse peak **b)** SDS-PAGE analysis of the peak fraction,as marked in panel a, that was used for crystallization experiments. M= MW marker, S= sample

The NBD of the transporters could be easily purified using Ni-IMAC followed by SEC yielding about 3 mg out of 200 ml bacterial culture (Figure A.16).

Unfortunately the NBD of YojI did not show any ATPase activity, but it could be bound to and eluted from ATP-linked agarose resin (Figure A.17).

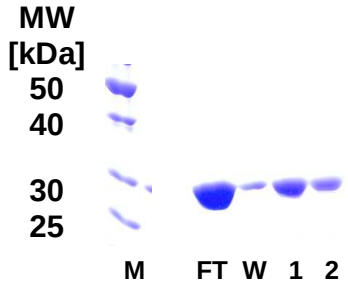


Figure A.17) Binding and elution of the Yoji NBD on ATP-linked agarose. M = marker, FT= flow-through, W = wash, 1 = elution 1, 2 = elution 2, the elution steps were done in presence of 20 mM ATP

Crystallization experiments were done at varying protein concentrations and buffer formulations as enlisted in Table A.7.

For the construct of the NBD of YojI, crystals were obtained in the two buffer conditions 0.1 M Tris pH 8, 0.15 M KSCN and 22 % PEG 3350 and Tris 0.1 M, pH 9 PEG 3350 18%, KSCN 0.15 M $CoCl_2$ in presence of Na_3VO_4, ADP, $MgCl_2$. These crystals are shown in figure 18. In both cases the crystals did not diffract when tested on the home source (CuKα radiation generated by a Nonius FR591 rotating anode generator).

Table A.7) List of the crystallization trials of the NBD of YojI. The concentration, additions prior crystal set ups, as well as the screens are indicated.

Protein	Concentrations	Additions prior to crystal set ups	Screens
YojI-NBD	5-20 mg/ml	$MgCl_2$	- Clear Strategy Screen™ I & II; pH 5.5, 6.5,7.5 and 8.5 (Molecular Dimensions) - GS002 - # 82009 Crystallization Basic Kit and #70437 Crystallization Extension Kit (Sigma Fluka)
YojI-NBD	5-20 mg/mlt	Na_3VO_4, ADP, $MgCl_2$	- Clear Strategy Screen™ I & II; pH 5.5, 6.5, 7.5, 8.5 (Molecular Dimensions) - GS002 - # 82009 Crystallization Basic Kit - # 70437 Crystallization Extension Kit (Sigma Fluka)
YojI-NBD	5-20 mg/ml	ATP, $MgCl_2$	- Clear Strategy Screen™ I & II; pH 5.5, 6.5, 7.5, 8.5 (Molecular Dimensions) - GS002 - # 82009 Crystallization Basic Kit (Sigma Fluka) - # 70437 Crystallization Extension Kit (Sigma Fluka)
YojI-NBD	10 mg/ml	$MgCl_2$	- GS041 (18% PEG 3350, 0.1M Tris (HOAc), 0.15 M KSCN pH 8, pH 8.5, pH 9)
YojI-NBD	10 mg/ml	Na_3VO_4, ADP, $MgCl_2$	-GS041 (18% PEG 3350, 0.1M Tris (HOAc), 0.15 M KSCN pH 8, pH 8.5, pH 9)

a) **b)**

Figure A.18) Crystals obtained of YojI-NBD **a)** in 0.1 M Tris pH 8, 0.15 M KSCN and 22 % PEG 3350 **b)** in 0.1 M Tris, pH 9, 18 % PEG 3350, 0.15 M KSCN and 0.01 M $CoCl_2$.

A.III Discussion

The crystallization of integral membrane proteins, such as ABC-transporters, is known to be a challenging task. A multitude of members of a protein family have to be screened in order to obtain stable and pure material for crystallization. This work focussed on half transporters having the NBD fused to the TMD, simplifying the cloning and purification of the proteins.
A purification protocol for the transporter YojI was established resulting in a stable protein that was used for crystallization experiments. Additionally, purification of the NBD of YojI yielded the protein in a monomeric state.

Although straightforward cloning techniques were exploited, the cloning of some transporters failed, as it was the case for the gene *mchF*; and several transporters could not be expressed at a reasonable amount, as it was the case for the transporter MclB. The problem was most likely due to the high toxicity of these transporters to the bacterial cell (Pons, Delalande et al. 2004). For the proteins that were expressing at detectable amounts, the expression level was dependent on the expression system and the expression conditions. For the proteins YojI and HlyB, the IPTG inducible vectors pET21/28 coding for a C-terminal His-tag were most suitable.

One crucial step in the purification process is the solubilization of the protein from the membrane. Detergent molecules have to replace the lipids of the membrane and shield the hydrophobic part of the protein. Moreover, the buffer conditions, such as pH and the ionic strength, need to be properly determined in order to avoid aggregation or denaturation of the protein. Since the detergent impairs biophysical characterization, it is not always easy to determine the folding state of a membrane protein. In this study, we have optimized the buffer and detergent condition for YojI.

We succeeded to purifiy this transporter in four different detergents at yields that allowed for setting up crystallization trials.

Although we obtained dimeric YojI as judged from SEC, no ATPase activity could be measured; therefore it is questionable, whether the protein YojI as purified here is in a functional state. The absent ATPase activity could be attributed to the lack of substrate addition, as the substrate (microcine J25) was unknown at the time the experiment was performed. For some ABC exporters such as the peptide tranporter TAP, the addition of the substrate is absolutely required to initiate ATP hydrolysis, i.e. ATP hydrolysis and substrate binding could also be tightly coupled in YojI (Chen, Abele et al. 2003) .

Future experiments could therefore include activity assays in the presence of the microcin J25 of YojI in either detergent solution, or reconstituted in a lipidic environment. Additionally, crystallization trials with the substrate microcin J25 could be performed.

Biochemical information on the NBDs of ABC transporters can be of interest to understand the functional mechanism of the entire transporter. The NBD of the transporter YojI was purified and could be bound to ATP-linked agarose indicating proper folding. Protein crystals were obtained but turned out to be non-diffracting. Based on results from bioinformatic predictions (Figure A.18) and partial tryptic digests (not shown) a future construct excluding the flexible linker regions located between the TMD and NBD and part of the C-terminus, could lead to a more stable protein domain with a better chance to provide diffracting crystals.

Based on the information obtained, screening further ABC transporter homologues for biochemical and biophysical characterization is worthwhile. In addition, specific binding partners such as DARPins and single chain antibody fragments could be raised against ABC transporters that can be purified, in order to increase the chance

to obtain well diffracting crystals (Ostermeier, Iwata et al. 1995; Sennhauser, Amstutz et al. 2007).

Figure A.18) Model of YojI-NBD, established by the software MODELLER in cartoon representation. The long unstructured N-terminus is clearly visible.

A.IV Material and Methods

A.IV.1 Cloning of the ABC-transporter

The plasmid encoding the ABC transporter *mclB* was a friendly gift of Sophie Sablé (Université La Rochelle, F).

The genes *mclB* and *mchF* derived from *E. coli* were amplified by PCR and cloned in the vectors pET21-Nt-His-tag, pET21-Ct-His-tag and pBAD24-Ct-His-tag. These vectors were already known to be suitable for these transporters (Frey 2007).

Cloning was done using PCR with primers for the desired protein, carrying the nucleotide sequence for the restriction site (BamHI and XhoI). The used primersare enlisted in Table A.8.

Table A.8) Primers used for cloning of the ABC-transporters.

Primer	Sequence
MchF_s1	TAGTCATGGATCCATGACTAACGGGAATTTCAGACAAATTATA
MchF_s2	TATGGTCGGTATCCGGGGAAC
MchF_as1	GTTCCCCGGATACCGACCATA
MchF_as2	TAGTCATCTCGAGAATAGAAATAATCCTGTCAACAGTTCTCAACG
MclB_as1	TAGTCATCTCGAGAATGGAAATAACTCTGTCAACAGTTTTCAAC
MchE_as1	TAGTCATCTCGAGGTCATTGAGCGGTCCTGTTGCACTGTGTT
MchE_pBAD_s	GCTAGTATTACTATGAGCTCATGTTTCGCCAGGATG
sp_hemo	CGGGATCCATGGATTCTTGTCATAAAAT
ap_hemo	CCGCTCGAGGTCTGACTGTAACTGATA
HlyBpBAD	CGCAGACCATGGATTCTTGTCATAAAAT

A.IV.2 Expression Tests

Expression of the proteins in the cloned vectors were tested in different expression systems, the cell lines BL21(DE3), C41(DE3) and C43(DE3) (Miroux and Walker 1996). For screening for the optimal expression conditions, pre-cultures of

transformed bacterial cells were used for inoculation of 5 ml liquid media. Bacterial cells were grown at 37 °C in the media LB or 2YT. Ampicillin and Kanamycin were used at a concentration of 0.1 mg/ml and 0.05 mg/ml, respectively. The protein over-expression was induced at an OD_{600} of 0.6. Multiple expression temperatures and times were tested: 37°C 4h, 30°C 4h and 18°C 20h, The arabinose-inducible vector was induced with 0.2%, 0.02%, 0.002% L-arabinose, the IPTG-inducible vecors were induced with 0.05, 0.1, 1 and 5 mM IPTG.

The overexpression was visualized by Dot Blot or Western Blot analysis, using antibodies specific to His-tagged protein (either a scFv ((Kaufmann, Lindner et al. 2002)) or commercial antibody (Quiagen)).

A.IV.3 Expression and Purification of the ABC-transporters

The full length transporter YojI and HlyB were expressed in shaking flasks with 500 ml of liquid media 2YT for protein purification. After induction with 0.1 mM IPTG the proteins were expressed at 18°C for 16 h.

After expression, cells were pelleted down and stored at -80°C for purification.

After thowing of the frozen pellets, the cells resuspended in 25 ml lysis buffer (20 mM Tris, 150 mM NaCl pH 7.5, DNAse and Complete Protease Inhibitor (Roche)) per pellet of 1 l bacterial culture and lysed using by French Pressing

The cellular debris was separated from the lysate with centrifugation at 5000 g for 15 min. Subsequently the membrane fraction was spun down by ultracentrifugation at 45 000 rpm for 1 h. The protein was solubilized in 30 ml of detergent-enriched buffer (20 mM Tris, 150 mM NaCl pH 7.5, respective detergent) per pellet. For the solubilization the detergents were used at the following concentrations: 1% DDM, 1 % Cymal 7, 2 % UDM and 5% Fos-12. During the purification, DDM was used at 0.03 %, Cymal 7 at 0.03 %, UDM at 0.1% and Fos-12 at 0.15%.

The non-solubilized material was spun down by a second ultracentrifugation step, before subjecting it to a pre-equilibrated Ni-NTA-IMAC column (Pharmacia).

Washing steps were performed with increasing imidazole concentrations via a 2-step gradient with 20 and 50 mM respectively. The protein was eluted in presence of high imidazole (250mM) and concentrated in a centricon (Millipore) concentrator with a molecular weight cut off of 100 kDa.

Size-exclusion chromatography was used in the buffer 20 mM Tris, 150 mM NaCl pH 7.5 or 20 mM Citrate, 150 mM NaCl, pH 6.5 for the final purification step and subsequently the samples were analysed by SDS-PAGE.

A.IV.4 Cloning, expression and purification of the NBDs

The NBD of the protein YojI was cloned using the restriction sites EcoRI and XhoI in the vector pET21-C-terminal His-tag. The used primers are enlisted in Table9. The NBD of the protein YojI was expressed in 2 x 100 ml of LB media for protein purification at 37°C for 4h after induction with 1mM of IPTG.

The cell pellet containing the YojI-NBD was lysed using a French Press using 20 ml of lysis buffer containing 20 mM Tris, 150 mM NaCl pH 7.5 and DNAse. The lysate was filtered with a pore size of 0.2 μm and subjected to a Ni-Protino column (Macherey & Nagel). Washing was done with 20 ml of Wash Buffer (20 mM Tris, 150 mM NaCl pH 7.5) and eluted with elution buffer (20 mM Tris, 150 mM NaCl pH 7.5, 250 mM imidazole). All purification steps were done at 4°C.

Table A.9) Primers used for the cloning of the NDBs of the ABC-transporters HlyB and YojI.

Primer	Sequence
HlyBNBD	GCCGGATCCGATATCTTTCGT
YojI_NBD_s	GCTATGAATTCACGGCGCAGGTGGC

A.IV.5 Screen on buffers and detergents

In order to screen for optimal purification conditions, different buffers and detergents were used. The protein was purified in 20 mM of the corresponding buffer, while the

NaCl was kept constant at 150 mM, the detergent was used at concentrations of 100X CMC for solubilization and 3X CMC for purification.

The following non-ionic detergents containing maltosides as uncharged, hydrophilic head groups were used: n-Dodecyl-β-Maltopyranoside (DDM) n-Undecyl-β-Maltopyranoside (UDM), n-Decyl-β-Maltopyranoside (DM) and n-Octyl-β-Maltopyranoside (OM). As non-ionic detergents containing glucosides as uncharged, hydrophobic head groups were used: n-Octyl-β-Glucopyranoside (OG), n-octyl-β-D-thioglucopyranoside (OTG). As lipid-like detergents n-Dodecylphosphocholine (Fos 12), n-Tetradecylphosphocholine (Fos 14), n-Hexadecylphosphocholine (Fos 16). Furthermore the cyclo-alkyl detergents 7-Cyclohexyl-1-hexyl-β-D-maltoside (Cymal 7), 6-Cyclohexyl-1-hexyl-β-D-maltoside (Cymal 6) were tested. As zwitterionic detergents N,N-Dimethyldodecylamine-N-oxide (LDAO), N,N-Dimethyldecylamine-N-oxide (DDAO) and 3-[(3-Cholamidopropyl)-dimethylammino]-1-propanesulfonate (CHAPS) were used. The protein YojI was first solubilized in DDM, subsequently the detergent was exchanged on the Ni-NTA column during the subsequent washing steps.

A.IV.6 Activity assays

The activity of the proteins was determined by detecting the quantity of phosphate liberated via formation of Malachite green – phosphomolybdate complex as decribed (Van Veldhoven and Mannaerts 1987). A standard curve was determined using a dilution series of H_3PO_4.

Binding assays were performed using ATP agarose (Sigma-Aldrich) in presence of 1 mM $MgCl_2$ by rotation at 4°C for 2h. The sample was washed 2 times using 20 mM Tris, 150 mM NaCl, pH 7.5 and eluted in presence of 20 mM ATP. For the full length transporter the buffers were supplemented with 0.03% DDM.

A.IV.7 Crystallization experiments

All crystallization experiments were done by sitting drop with vapour diffusion against a reservoir of varying buffer conditions. Different crystallization screens based on different buffer compositions were done at the NCCR crystallization facility at 4°C and 20°C. The crystal screens used are listed in the appendix.

A. V Literature

Azpiroz, M. F., E. Rodriguez, et al. (2001). "The structure, function, and origin of the microcin H47 ATP-binding cassette exporter indicate its relatedness to that of colicin V." Antimicrob Agents Chemother **45**(3): 969-72.

Blond, A., M. Cheminant, et al. (2001). "Solution structure of microcin J25, the single macrocyclic antimicrobial peptide from Escherichia coli." Eur J Biochem **268**(7): 2124-33.

Borbat, P. P., K. Surendhran, et al. (2007). "Conformational motion of the ABC transporter MsbA induced by ATP hydrolysis." PLoS Biol **5**(10): e271.

Chen, M., R. Abele, et al. (2003). "Peptides induce ATP hydrolysis at both subunits of the transporter associated with antigen processing." J Biol Chem **278**(32): 29686-92.

Dawson, R. J. and K. P. Locher (2006). "Structure of a bacterial multidrug ABC transporter." Nature **443**(7108): 180-5.

Dean, M. and R. Allikmets (1995). "Evolution of ATP-binding cassette transporter genes." Curr Opin Genet Dev **5**(6): 779-85.

Dean, M., A. Rzhetsky, et al. (2001). "The human ATP-binding cassette (ABC) transporter superfamily." Genome Res **11**(7): 1156-66.

Delgado, M. A., P. A. Vincent, et al. (2005). "YojI of Escherichia coli functions as a microcin J25 efflux pump." J Bacteriol **187**(10): 3465-70.

Frey, D. (2007). Biochemical and Structural Studies of Dimethylarginine Dimethylaminohydrolase-1, the ABC Transporter Superfamily and the Sodium Citrate Transporter CitS. MNF Zurich, University of Zurich. **PhD:** 204.

Havarstein, L. S., D. B. Diep, et al. (1995). "A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export." Mol Microbiol **16**(2): 229-40.

Higgins, C. (1992). "ABC transporters: from microorganisms to man." Annu Rev Cell Biol **8**: 67-113.

Higgins, C. F. and K. J. Linton (2004). "The ATP switch model for ABC transporters." Nat Struct Mol Biol **11**(10): 918-26.

Holland, I. B. and M. A. Blight (1999). "ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans." J Mol Biol **293**(2): 381-99.

Hopfner, K. P., A. Karcher, et al. (2000). "Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily." Cell **101**(7): 789-800.

Hwang, J., X. Zhong, et al. (1997). "Interactions of dedicated export membrane proteins of the colicin V secretion system: CvaA, a member of the membrane fusion protein family, interacts with CvaB and TolC." J Bacteriol **179**(20): 6264-70.

Kaufmann, M., P. Lindner, et al. (2002). "Crystal structure of the anti-His tag antibody 3D5 single-chain fragment complexed to its antigen." J Mol Biol **318**(1): 135-47.

Kobayashi, N., K. Nishino, et al. (2001). "Novel macrolide-specific ABC-type efflux transporter in Escherichia coli." J Bacteriol **183**(19): 5639-44.

Locher, K. P. (2009). "Review. Structure and mechanism of ATP-binding cassette transporters." Philos Trans R Soc Lond B Biol Sci **364**(1514): 239-45.

Locher, K. P., A. T. Lee, et al. (2002). "The E. coli BtuCD structure: a framework for ABC transporter architecture and mechanism." Science **296**(5570): 1091-8.

Miroux, B. and J. E. Walker (1996). "Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels." J Mol Biol **260**(3): 289-98.

Muller, K. M., C. Ebensperger, et al. (1994). "Nucleotide binding to the hydrophilic C-terminal domain of the transporter associated with antigen processing (TAP)." J Biol Chem **269**(19): 14032-7.

Oldham, M. L., D. Khare, et al. (2007). "Crystal structure of a catalytic intermediate of the maltose transporter." Nature **450**(7169): 515-21.

Ostermeier, C., S. Iwata, et al. (1995). "Fv fragment-mediated crystallization of the membrane protein bacterial cytochrome c oxidase." Nat Struct Biol **2**(10): 842-6.

Pons, A. M., F. Delalande, et al. (2004). "Genetic analysis and complete primary structure of microcin L." Antimicrob Agents Chemother **48**(2): 505-13.

Pons, A. M., I. Lanneluc, et al. (2002). "New developments in non-post translationally modified microcins." Biochimie **84**(5-6): 531-7.

Quinton, P. M. (1999). "Physiological basis of cystic fibrosis: a historical perspective." Physiol Rev **79**(1 Suppl): S3-S22.

Rees, D. C., E. Johnson, et al. (2009). "ABC transporters: the power to change." Nat Rev Mol Cell Biol **10**(3): 218-27.

Rosenberg, M. F., R. Callaghan, et al. (1997). "Structure of the multidrug resistance P-glycoprotein to 2.5 nm resolution determined by electron microscopy and image analysis." J Biol Chem **272**(16): 10685-94.

Sauna, Z. E. and S. V. Ambudkar (2007). "About a switch: how P-glycoprotein (ABCB1) harnesses the energy of ATP binding and hydrolysis to do mechanical work." Mol Cancer Ther **6**(1): 13-23.

Seeger, M. A. and H. W. van Veen (2009). "Molecular basis of multidrug transport by ABC transporters." Biochim Biophys Acta **1794**(5): 725-37.

Sennhauser, G. (2004). Expression and purification of an ABC transporter. Biochemische Insitut. Zuerich, University of Zuerich. **Diplom**.

Sennhauser, G., P. Amstutz, et al. (2007). "Drug export pathway of multidrug exporter AcrB revealed by DARPin inhibitors." PLoS Biol **5**(1): e7.

Story, R. M. and T. A. Steitz (1992). "Structure of the recA protein-ADP complex." Nature **355**(6358): 374-6.

Urbatsch, I. L., B. Sankaran, et al. (1995). "Both P-glycoprotein nucleotide-binding sites are catalytically active." J Biol Chem **270**(45): 26956-61.

Van Veldhoven, P. P. and G. P. Mannaerts (1987). "Inorganic and organic phosphate measurements in the nanomolar range." Anal Biochem **161**(1): 45-8.

Vincent, P. A. and R. D. Morero (2009). "The structure and biological aspects of peptide antibiotic microcin J25." Curr Med Chem **16**(5): 538-49.

Walker, J. E., M. Saraste, et al. (1982). "Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold." Embo J **1**(8): 945-51.

Ward, A., C. L. Reyes, et al. (2007). "Flexibility in the ABC transporter MsbA: Alternating access with a twist." Proc Natl Acad Sci U S A **104**(48): 19005-10.

Chapter B -Biophysical characterization of SGT1

B.I Introduction

B.I.1 Inflammation and innate immunity

The word “inflammation” originates from the Latin word “inflammare”, meaning “to set on fire”. Inflammation in an affected area is generally characterized by five clinical signs: redness and increase of temperature, swelling, loss of function and pain. These macroscopic observations are a result of a set of reactions in the mammalian immune system that are triggered in response to foreign objects in the body, such as bacteria, viruses but also in response to small molecules such as ATP.

In inflammatory processes, immune cells are recruited to the site of inflammation, pathogens are removed by macrophages and neutrophils, and the adaptative immune system and the complement cascade are activated.

The inflammasome plays a central role for this multitude of reactions (Martinon, Burns et al. 2002).

A hallmark during the inflammatory reaction is the formation of the NLRP3 (NACHT-, LRR- & Pyrin-domain containing) inflammasome, which activates the inflammatory caspase-1 (Martinon, Petrilli et al. 2006). Activated caspase-1 then processes the mediators of the inflammatory response, e.g. IL-1β (Martinon, Burns et al. 2002). The protein SGT1, the focus of this project, plays an important role in the formation of the inflammasome.

B.I.2 The inflammatory cascade

Inflammatory cascades are initiated by pathogen-recognition receptors (PRRs) that recognize pathogen-associated molecular patterns (PAMPs) thereby triggering a

response to the microbial and endogenous stimuli. PAMPs can be lipopolysaccharides (LPS) or peptidoglycans (PGNs) derived from the degradation of bacterial cell wall, viral DNA, but also ATP, DNA, RNA and uric acid.

The PRRs can be divided into two groups; the Toll-like receptors (TLRs), which are transmembrane receptors sensing the inflammatory stimuli already outside the cell, and the nucleotide binding domain, leucine rich repeat containing receptors (NLRs), which are intracellular receptors. They are responsible for the pathogen recognition inside the cell. Both, TLRs and NLRs are constituted of several domains and are binding to host-derived molecules. Their common domain is the leucine rich repeat (LRR) domain, responsible for the detection of the pathogen.

The family of NLRs comprises the NLRA (member: CIITA class II, major histocompatibility complex, transactivator), NLRPs (NLR family, pyrin domain containing proteins), the NLRB (member: NAIP NLR family, apoptosis inhibitory protein), the, NLRX (NOD-like receptor with "unknown" domain, members NOD1&2 nucleotide-binding oligomerization domain containing) and NLRC (NLR forming, CARD domain containing). Next to the C-terminal LRR-domain, all NLRs share also a NACHT domain (domain present in NAIP, CIITA, HET-E and TP1) (Koonin and Aravind 2000), which was proposed to be a switch for conformational changes and has an ATPase activity. It is also responsible for the oligomerization of NLRs.

The N-terminal recruitment domain of NLRs is variable and can be a CARD (caspase recruitment domain), BIR (Baculoviral inhibitor of apoptosis Repeats), PYD (pyrin) or a transmembrane domain (Ye and Ting 2008). These domains are responsible for inter-protein interactions and confer specificity. Different stimuli activate the NLRs, NLRP3, NOD1 and NOD2 for example are activated by mesodiamiopimelic acid and muramyl dipeptide, the degradation products of PGN, (Martinon, Agostini et al. 2004; Fritz, Ferrero et al. 2006).

Different stimuli activate the NLRs, which in turn activate different inflammasomes and different responses.

NLRP3, NOD1 and NOD2 for example are activated by meso-diaminopimelic acid and muranyldipeptide (MDP), both degradation products of PGNs. Upon activation, NOD2 forms a complex with RIP2 (receptor interacting kinase 2) that activates NF-κB and Pro-IL-1β synthesis (Meylan and Tschopp 2005). NLRP3 activation induces inflammasome formation, which activates caspase-1 that processes ProIL-1β to functional IL-1β (Figure B.1).

The inflammatory response is highly regulated by pro- and anti-inflammatory cytokines. A deregulation might lead to immunodeficiency or auto-inflammatory diseases (Alheim and Bartfai 1998). Also mutations in the inflammatory proteins can lead to diseases. Some well-studied examples are mutations in NLRP3, which lead to familial-cold auto-inflammatory syndrome Muckle-Wells syndrome (Aganna, Martinon et al. 2002; Dode, Le Du et al. 2002).

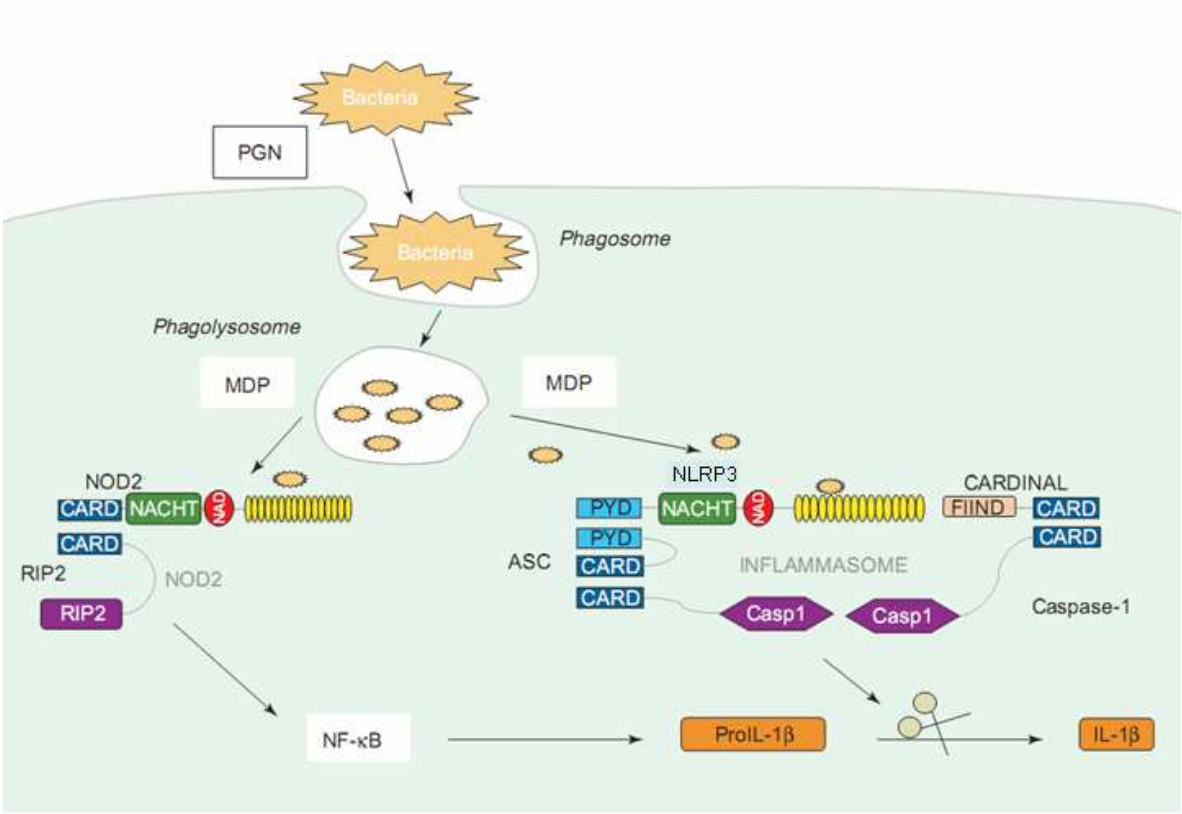


Figure B.1) Model for the activation of the inflammasome by PGN derivates from (Martinon and Tschopp 2005). After phagocytosis of bacteria, MDP, the degradation product of PGN activates NOD2. A NOD2-RIP2 complex is formed. The resulting RIP2 activation leads to NF-κB activation. ProIL-1β is synthesized. The release of MDP also triggers the activation of NLRP3. The inflammasome complex (NLR3, cardinal, ASC and caspase-1) is assembled and activates caspase-1. ProIL-1β is then converted to IL-1β.

B.I.3 The inflammasomes

Three types of inflammasomes have so far been described: the NLRP1, NLRP3 and the NLRP4 inflammasomes (Figure B.2) (Martinon, Gaide et al. 2007). NLRP1 and NLRP3 serve as the scaffolding proteins for these multi protein assemblies.

NLRP3 interacts with the adaptor protein ASC (apoptosis-associated speck-like protein) which is comprised of a CARD and a PYD, via PYD-PYD interaction. ASC is then linked with the help of CARD-CARD interactions to caspase-1. The NLRP1 inflammasome has a similar architecture, ASC links NLRP1 to caspase-1 and caspase-5

(Martinon, Burns et al. 2002; Agostini, Martinon et al. 2004). In the NLRP4 inflammasome, no adaptor protein has been found. NLRP4 interacts directly with caspase-1 via the CARD domain (Poyet, Srinivasula et al. 2001).

Caspases that bind to the inflammasomes are activated and are ready to process cytokines such as proIL-1β to its active form IL-1β, which is the main signalling molecule in the inflammatory response.

B.I.4 Inflammatory caspases

Inflammatory caspases are activated in the innate immune response as described above (Martinon and Tschopp 2004). They belong to the familiy of caspases, which are endopeptidases. This protein family is named after their catalytic residue, *c*ysteine, involved in cleaving peptide bonds C-terminal from *asp*artate.

Human inflammatory caspases are caspase-1, -4, -5 and -12. They are constituted of a pro-domain with a caspase-recruitment domain (CARD) followed by a catalytic domain.

Caspase-1 (interleukin-1β converting enzyme, ICE) was the first caspase to be identified; it activates the inflammatory ProIL-1β to its mature active form IL-1β.

Little is known with respects to the inflammatory pathways of caspases-4, -5 and -12. Caspase-5 is known to cleave caspase-3 (Schwartz, Yamamoto et al. 1999) and to process the transcription factor Max in vitro (Krippner-Heidenreich, Talanian et al. 2001). Caspase-5 together with caspase-1 were found to be components of the NLRP1 inflammasome, a complex involved in their reviewed by (Martinon, Burns et al. 2002).

For caspase-4 and caspase-12 it has been show, that they are involved in the so called endoplasmatic stress response (Nadiri, Wolinski et al. 2006); a mechanism acting against the viral infection by inducing protein overproduction and aggregation.

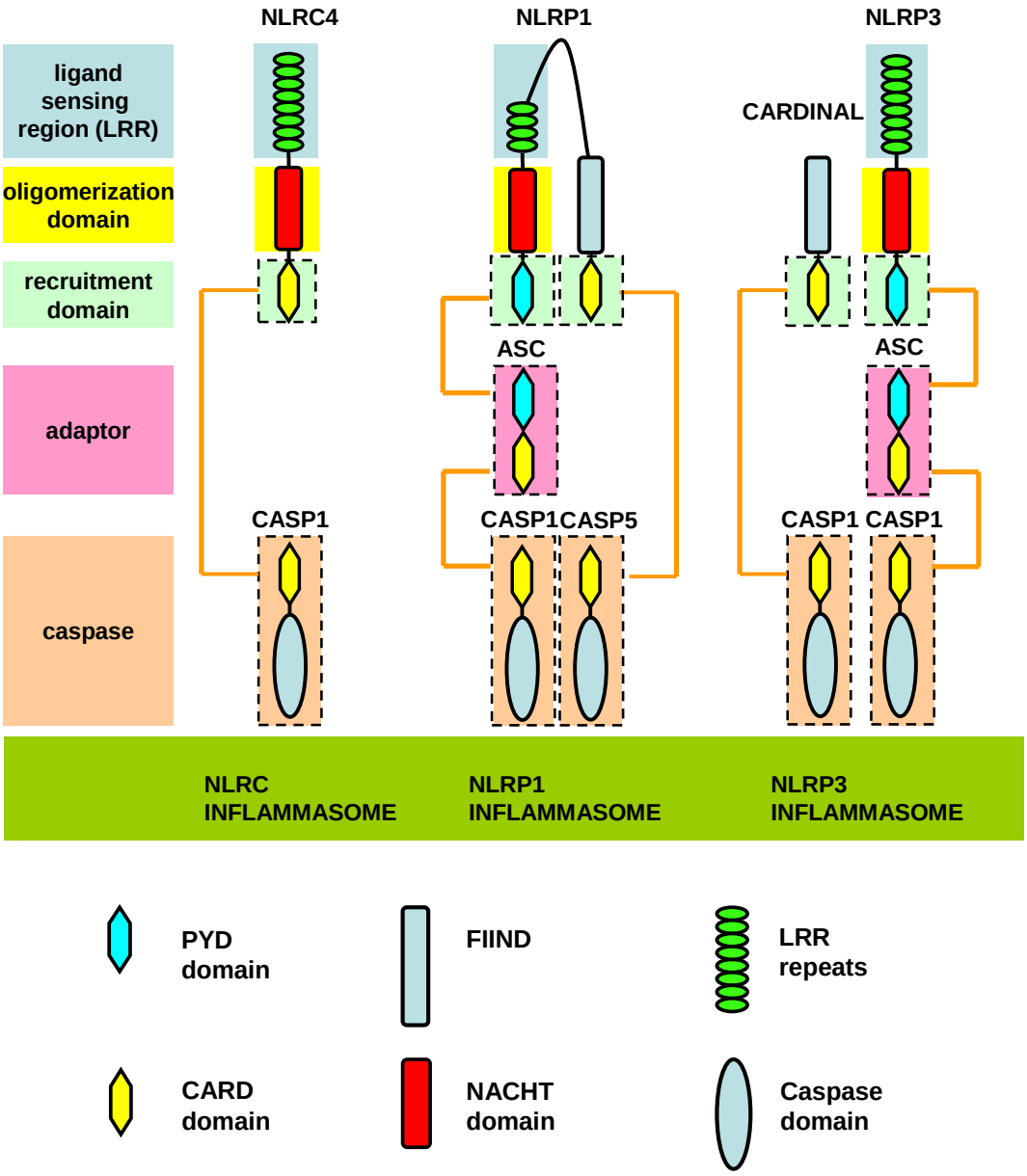


Figure B.2) Schematic overview of the interactions mediated by the scaffolding proteins of the NLR family for creating the different inflammasones (NLRC4-, NLRP1- and NLRP3-inflammasome, respectively). The oligomerization is initiated by the LRRs of NLRC4, NLRP1 and NLRP3 after binding to a ligand. The protein-protein interactions are mediated by homotypic domain dimerization. The scaffolding protein interacts with the inflammatory caspases either via a scaffolding protein or directly via the recruitment domain. The activated caspase is then processing the cytokine proIL-1β to IL-1β, starting the inflammatory response. Picture modified after (Martinon and Tschopp 2007).

B.I.5 The role of SGT1 in inflammation

For the recruitment of NLRP3 to the inflammasome, HSP90 (Heat shock protein 90) and SGT1 are important. HSP90 is a chaperone well-known for being involved in protein folding and cell signalling (Lee, Jacob et al. 2004). SGT1 is a co-chaperone and "client adaptor". It recruits specific clients to the HSP90 co-chaperone complex (Mayor A. et al (2007)).

In the absence of HSP90 NLRP3 is unstable and is degraded by the proteasome. There is a possibilty that unbound NLRP3 might even autoactivate resulting in inflammasome formation. Binding of NLRP3 to HSP90 and SGT1 is observed and interpreted as stabilization and inactivation of NLRP3. Upon activation, SGT1 and HSP90 are released and NLRP3 adopts an alternative conformation, allowing it to interact with other inflammasome components, as depicted in Figure B.3 (Mayor, Martinon et al. 2007).

The activation mechanism of other members of the NLR family is unknown, but it is presumed to be similar (Mayor, Martinon et al. 2007). Additionally, it has been shown, that SGT1 is also involved in the pathways mediated by NOD1, NOD2, NLR1-4, IPAF and NAIP (da Silva Correia, Miranda et al. 2007; Mayor, Martinon et al. 2007).

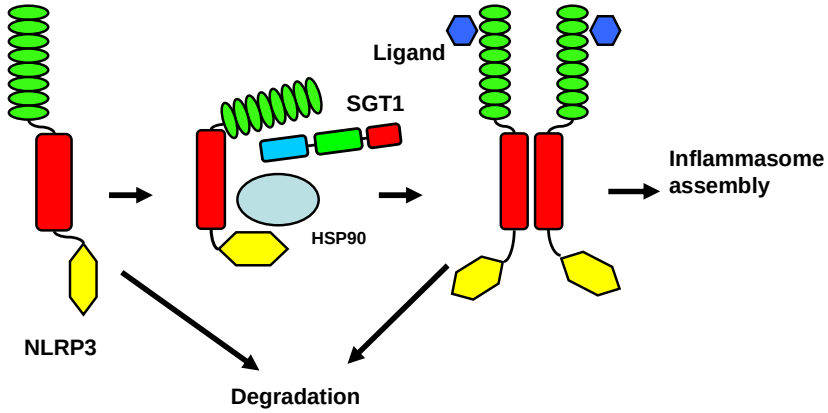


Figure B.3) Proposed model for SGT1 acting as a cochaperone for HSP90 in the inflammatory response after Mayor et al (2007). SGT1 and HSP90 are keeping NLRP3 in an active, but signalling competent state. As soon as a ligand is bound to the LRR, NLRP3 multimerizes and the other inflammasome components are recruited. NLRP3 alone is supposed to be instable and degraded by the proteasome. Picture modified after (Mayor, Martinon et al. 2007).

B.I.6 SGT1 - function and its architecture

Figure B.4) The architecture of the human protein SGT1. This protein is constituted of the TPR domain, the CS domain and the SGS domain. The TPR domain is involved in oligomeric assembly and binding to skp1 during mitosis, the CS domain mediates the binding to HSP90 & RAR1 and the SGS domain binds to calcyclin and other S100 proteins. These domains are linked by variable regions (VR 1 & 2).

SGT1 is a highly conserved eukaryotic (Shirasu and Schulze-Lefert 2003). Along with its function in the mammalian immune system (Mayor, Martinon et al. 2007), and the hypersensitivity response in plants, it is also involved in cell cycle regulation, where it has been first discovered in yeast. It was subsequently named SGT1, which stands for suppressor of the G2 allele of skp1. It has been shown that the two proteins, SGT1 as

well as skp1 are required for the G1/S and G2/M transition in the cell cycle (Kitagawa, Skowyra et al. 1999).

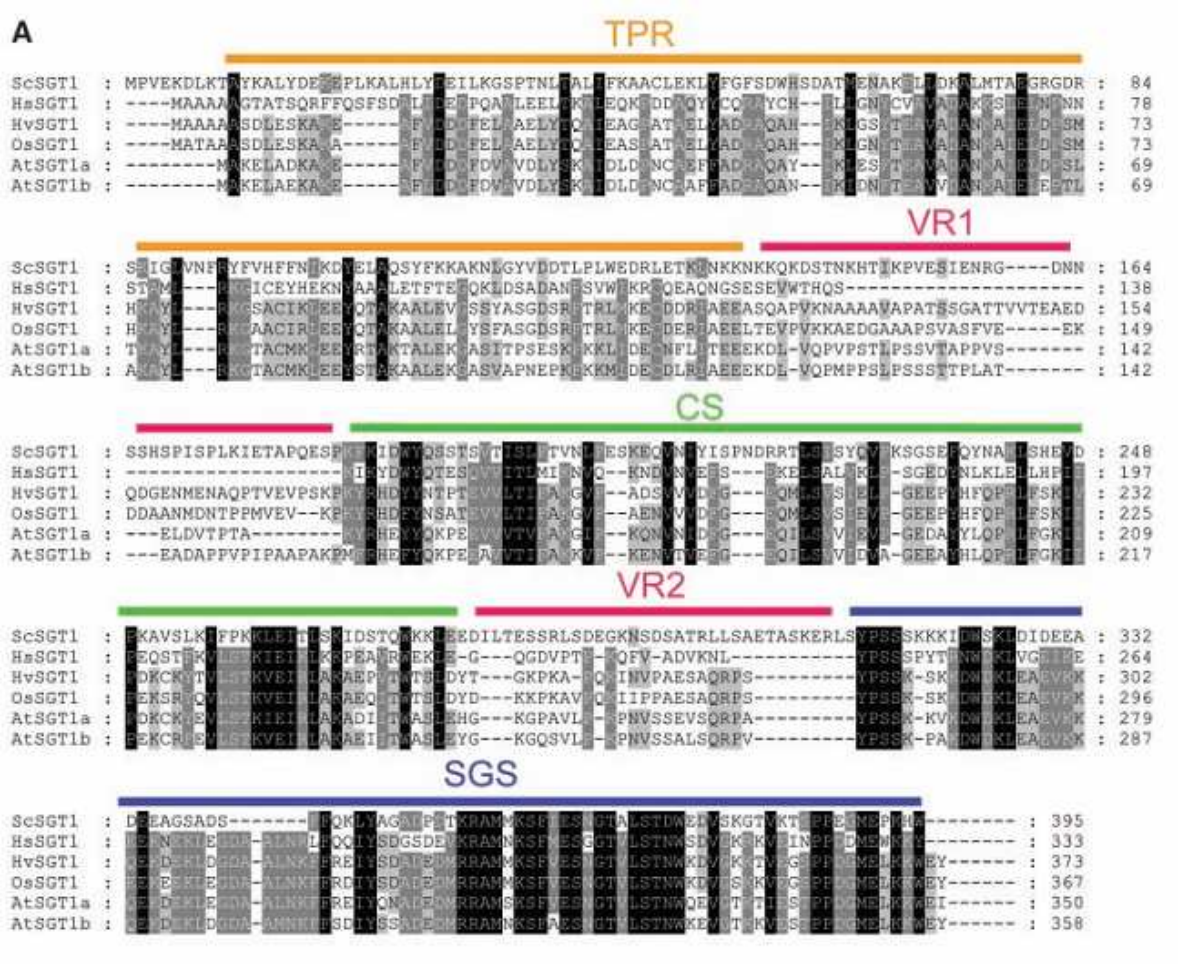


Figure B.5) Sequence alignment of SGT1 from different organism. Sc= *Saccharomyces cerevisiae,* Hs=human, Hv= barley, Os=rice, At= *Arabidopsis thaliana.* The coloring indicates the level amino acid identity or similarity: black= 100%,dark grey = 80%, light grey = 60 %; taken from (Azevedo, Sadanandom et al. 2002).

In human, two isoforms of SGT1 are found, which differ in 33 amino acids in the TPR domain. The function and the architecture are the same. SGT1 consists of three domains, as depicted in Figure B.4. The tetratrico peptide repeat (TPR) domain is located at the N-terminus, followed by a central CS domain (*C*HORD and *S*GT1 motif or also called p23-like domain) and at a C-terminal a SGT1 specific (SGS) domain. Variable regions (VR1 and VR2) are found between these domains, when comparing

sequences of different species (Figure B.5). NMR experiments showed that these parts of the protein have no defined structure (Lee, Jacob et al. 2004).
Homologous of human SGT1 have been found in many species. Most of them share this three domain architecture, but in the nematodes *Caenorhabditis* and *Brugia* the protein SGT1 just comprises two domains, the CS-domain and the SGS-domain (reviewed by (Shirasu 2008).

Structural information of SGT1 of individual domains is available.
The TPR domain is a protein-protein interaction domain, which is binding to skp1 in kinetochore assembly during mitosis (Catlett and Kaplan 2006). Besides, recent studies have shown that the TPR domain is mainly responsible for the stability of the SGT1 protein in plants, probably by forming oligomers (Azevedo, Betsuyaku et al. 2006; Nyarko, Mosbahi et al. 2007). X-ray structures of different TPR domains have been determined (Main, Xiong et al. 2003). They show a mainly alpha helical architecture.

The CS domain plays a central role in plant disease resistance. Two binding sites located on opposite sites of the domains allow the simultaneous interactions with RAR1 and the N-terminal domain of HSP90 (Boter, Amigues et al. 2007). whereas RAR1 shares with the CS domain partially the same binding site on HSP90 (Kadota, Amigues et al. 2008). Although RAR1 is dispensable in disease resistance, it enhances the function of SGT1 on HSP90, which is then able to act as a recruiter for chaperone activities on multiprotein complexes (Boter, Amigues et al. 2007). The NMR structure of the CS domain of human SGT1, showed that the domain adopts a 2 x 2 and 1 x 3 antiparallel ß-strand architecture (Lee, Jacob et al. 2004)). The topology is related to the one of the protein p23, which also shares high sequence homology with the CS

domain and also interacts with the N-terminal domain of HSP90 (Ali, Roe et al. 2006) (Figure B.6).

In mammalian immune response, the precise interactions are still not determined. So far it is only known that the presence of the CS and SGS domain of SGT1 is essential for the interaction of NLRP3 to HSP90 (Mayor, Martinon et al. 2007).

Recently, the crystal structure of the complex formed between the N-terminus of Hsp90 from *Hordeum vulgare* (70 % identical to human) and the CS domain of Sgt1a from *Arabidopsis thaliana* (36 % identical to human) shows how SGT1 interacts with HSP90 (Figure B.6).

Most of the interacting residues of the CS domain are located on the surface of the four stranded β-sheet. This explains the binding of this domain to HSP90 in absence of ATP, conformational changes of HSP90 during the ATPase cycle exacerbate the binding (Boter, Amigues et al. 2007; Kadota, Amigues et al. 2008; Zhang, Boter et al. 2008).

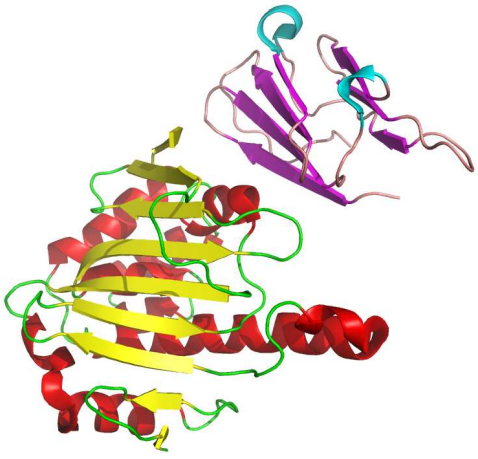


Figure B.6) X-ray structure of the CS domain of SGT1 of *Arabidopsis thaliana* and the N-terminal domain of *Hordeum vulgare* HSP90 (2JKI.pdb). The interactions are mediated by the 2x2 stranded β-sheet. The colour coding is analogous to a), for HSP90 α-helices are coloured in red, β-sheets are coloured in yellow.

The human SGS domain can interact with calcyclin and other S100 proteins in a calcium specific manner (Nowotny, Spiechowicz et al. 2003). Currently its purpose in the pathogen response reaction is unknown, but it is assumed that it serves as an interaction domain for some LRR- containing proteins, like Mla1 from plant and NOD1 from mammals (Dubacq, Guerois et al. 2002; Bieri, Mauch et al. 2004; da Silva Correia, Miranda et al. 2007). NMR and CD spectroscopy studies of the SGS domain indicated that it is not a well folded globular domain having only a limited degree of α-helical secondary structure.

B.I.7 Short summary of Results

In this part of the PhD thesis the expression, purification and biophysical characterization of the human SGT1 protein is described. The obtained results will help in future attemps to characterize hSGT1 structurally and functionally.

Different constructs of human SGT1 and its domains were designed. The oligomeric state of full length SGT1 and constructs comprising the first domains thereof (TPR-CS$_{D1-257}$ and TPR-CS$_{T14-235}$) were analysed by multi angle light scattering in the absence of cofactors or binding partners.

To estimate the content of secondary structure, circular dichroism spectra were recorded. They revealed that SGT1 has characteristics of a highly alpha helical protein, presumably mediated by the TPR domain, as indicated by the spectra of TPR-CS$_{D1-257}$ and TPR-CS$_{T14-235}$.

Fluorimetrical studies by recording spectra in presence and in absence of an denaturing agent, showed two different states for the full length SGT1 folding, whereas no clear discrimination of folding states for the SGS domain was detected.

Indeed, ^{1}D NMR spectroscopy confirmed that the SGS domain does not show any secondary structure.

Crystallization experiments of the full length protein SGT1 and on the constructs comprising the first two domains TPR-CS$_{T14-235}$ and TPR-CS$_{D1-257}$, were performed. Although, both Thermofluor-based high-throughput buffer formula optimization, and hand set precipitation screens were used to optimize the crystallization conditions, no diffracting crystals were obtained.

B.II Results of SGT1:

B.II.1 Constructs for biophysical characterization and crystallization trials

Four constructs of the human protein SGT1b (NM_006704), one of the two isoforms found in humans, were used for the subsequent work: the full length SGT1, the SGS domain, TPR-CS$_{D1-257}$ and TPR-CS$_{T14-235}$ were expressed as GST fusion proteins, depicted in Figure B.7. The constructs TPR-CS$_{D1-257}$ and TPR-CS$_{T14-235}$ comprised the TPR and CS domains only and were determined by limited proteolysis from the full length protein, as described later.

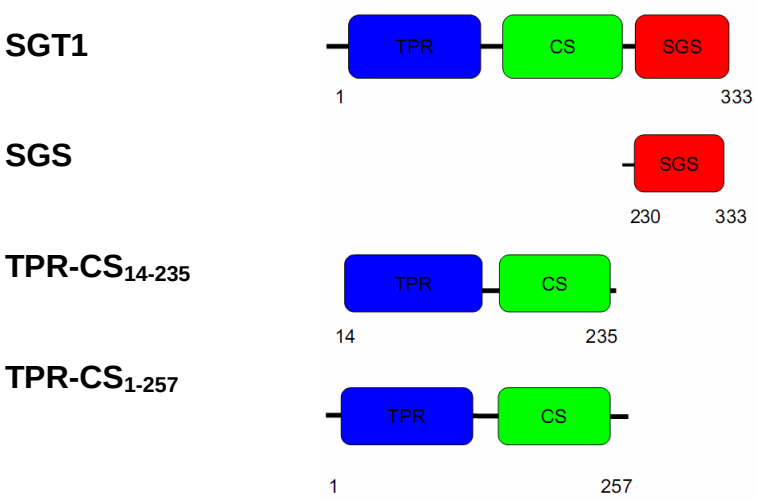


Figure B.7) The constructs of SGT1 and its domains used in this PhD thesis. The full length construct comprises amino acids 1-333 and the SGS domain the amino acids 230-333. The constructs determined by degradation experiments TPR-CS$_{14-235}$ and TPR-CS$_{1-257}$ comprise the TPR and the CS-domain only.

B.II.2 Expression and Purification of SGT1 and its domain SGS

Optimal expression was observed for the fusion protein GST-SGT1 and for the fusion protein GST-SGS at 37°C for 4 h after induction in LB media.

The full length protein SGT1 was co-purified with cleavage products after ion exchange chromatography, as seen in Figure B.8 a and b and confirmed by previous mass spectrometric analysis by G. Hausammann, revealing degradation products with molecular weights of 38.6 kDa, 35.6 kDa, 32.2 kDa and 25.7 kDa determined by mass spectroscopy (MALDI) (Hausammann 2007). In contrast to this, no degradation products

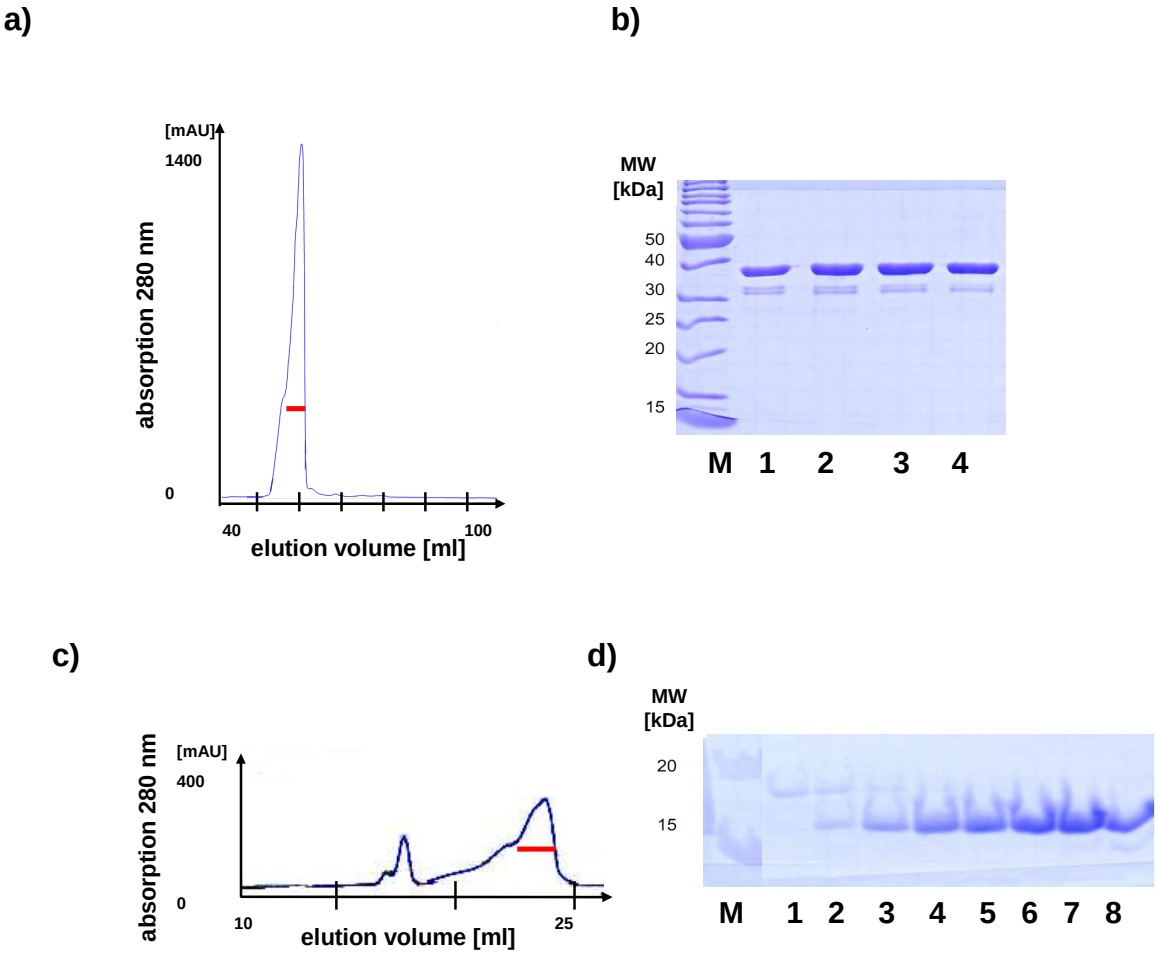


Figure B.8) Ion exchange chromatography and SDS-PAGE analysis of the different constructs of SGT1 **a)** Profile of the cation-exchange chromatography of SGT1 with a gradient of NaCl from 0 to 250 mM; **b)** SDS-PAGE analysis of the peak fractions, as marked in a); M= Molecular weight marker. **c)** Elution profile of anion-exchange chromatography of the SGS domain of SGT1. **d)** SDS-PAGE analysis of the peak fractions; M= Molecular weight marker,as marked in c).

were occurring when purifing the SGS domain alone, as seen in SDS-PAGE electrophoresis of the fractions eluted from ion exchange chromatography .

(Figure B.8 a-d). ESI-MS confirmed the identity of the protein by measuring a MW of 38.6 kDa for the full length construct SGT1 and 12.6 kDa for the construct SGS.

B.II.3 Determination of stable fragments

It is known from the literature that purified SGT1 is prone to degradation (Catlett and Kaplan 2006). Successful crystallization of the full length construct could be impeded by this behaviour. Limited proteolysis experiments were carried out for the design of new constructs being resistant to degradation. Subsequently these constructs were cloned; they

contained the first two domains the TPR and the CS domain. These constructs were named TPR-CS$_{1-257}$ and TPR-CS$_{14-235}$. The limited proteolysis experiments were performed using two different approaches. In the first one, the degradation of the protein was observed at room temperature during a time course of 10 days. The SDS-PAGE already revealed a stable fragment with a molecular weight between 25-30 kDa, as seen in Figure B.9a. ESI-MS in collaboration with Dr. S. Chesnov (FGCZ) subsequently identified this most stable fragment, having a mass of 29096.2 Da. As calculated with the program GPMAW (Peri, Steen et al. 2001), this molecular weight refers to an N-terminal degradation product, being composed of the TPR and the CS domain only and comprises the amino acids 1-257. It matches the prediction by the sequence aligment and domain assignment from Azevedo and coworkers (Azevedo, Sadanandom et al. 2002), showing that this fragment lacks the SGS domain.

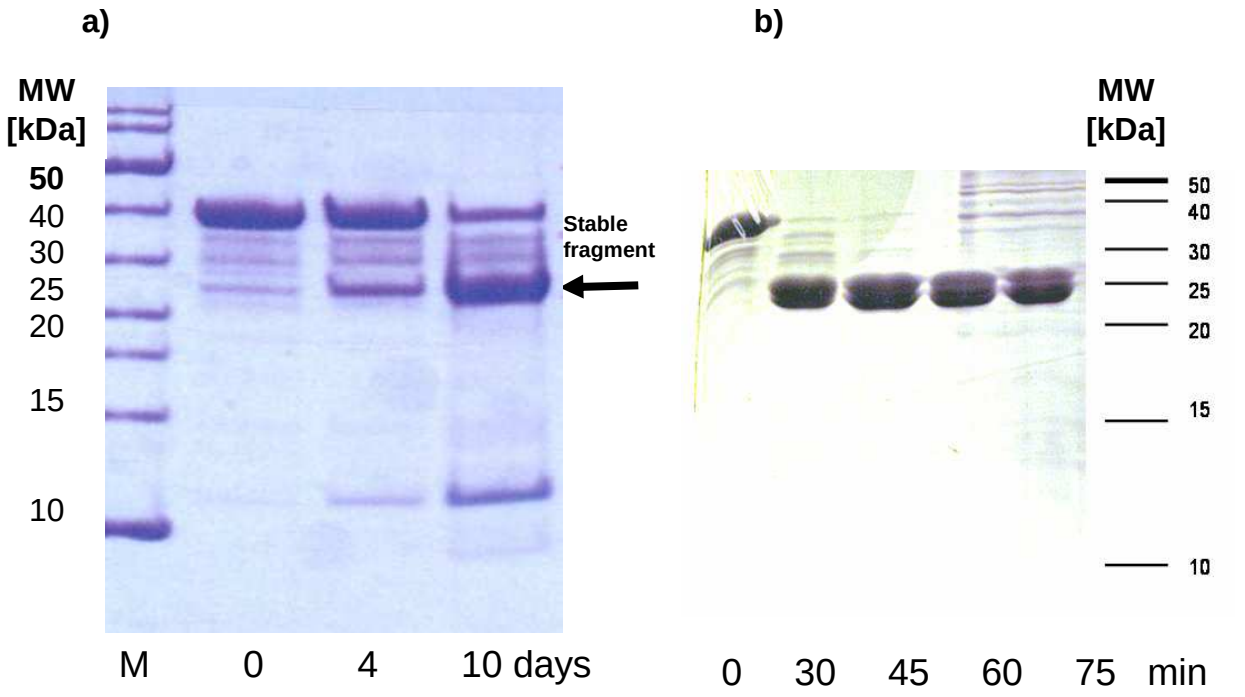


Figure B.9) **a)** Determination of proteolysis resistant constructs, by keeping the protein SGT1 for 10 days at room temperature. SDS-PAGE analysis of the protein directly, (0) after 4 and after 10 days. M= Molecular weight marker. **b)** Determination of stable constructs by partial tryptic digest. SGT1 was incubated with trypsin, the reaction was stopped after 30, 45, 60 and 75 minutes by precipitation with TCA. M= Molecular weight marker.

The subsequently cloned construct TPR-CS$_{1-257}$ includes amino acids 1- 257, having a theoretical molecular weight of 29095.8 Da (determined by protparam), without respecting the additional amino acids being introduced by the expression system.

The second approach was to use a limited tryptic digestion for the determination of the stable protein boundaries. Two main fragments were obtained (Figure B.9b). The protein SGT1 was incubated with trypsin and the proteolysis reaction was stopped by TCA precipitation at different time points. MALDI analysis revealed two degradation products with masses of 25342.0 and 28420.0 Da. Additionally N-terminal sequencing in colaboration with Dr. R. Brunisholz (FGCZ) identified one stable fragment starting from the amino acid 14 to 235. Only this fragment was consistent with the N-terminal sequence and the mass. The subsequently cloned construct was named TPR-CS$_{14-235}$, having a theoretical molecular weight of 25320 Da (determined by protparam), which is in the error range of the method (MALDI). The C-terminal end of the fragment is located after the CS domain, in the beginning of the variable region 2 (VR2) at the amino acid 235.

B.II.4 Expression and Purification of the constructs TPR-CS$_{14-235}$ and TPR-CS$_{1-257}$

The constructs TPR-CS$_{14-235}$ and TPR-CS$_{1-257}$ were expressed in the vector pGEX6-P1, possessing an N-terminal GST-tag. Best expression were observed for the fusion proteins GST-TPR-CS$_{T14-235}$ and GST-TPR-CS$_{D1-257}$ in *E. coli* at 20°C for 18 h after induction in LB media.

The proteins TPR-CS$_{T14-235}$ and TPR-CS$_{D1-257}$ were purified without degradation products, as seen in SDS-PAGE electrophoresis of the fractions eluted from ion exchange chromatography (Figure B.10 a-d). ESI-MS confirmed the identity of the proteins by measuring a MW of 26238 kDa for the construct TPR-CS$_{14-235}$ and 29126 Da for the construct TPR-CS$_{1-257}$.

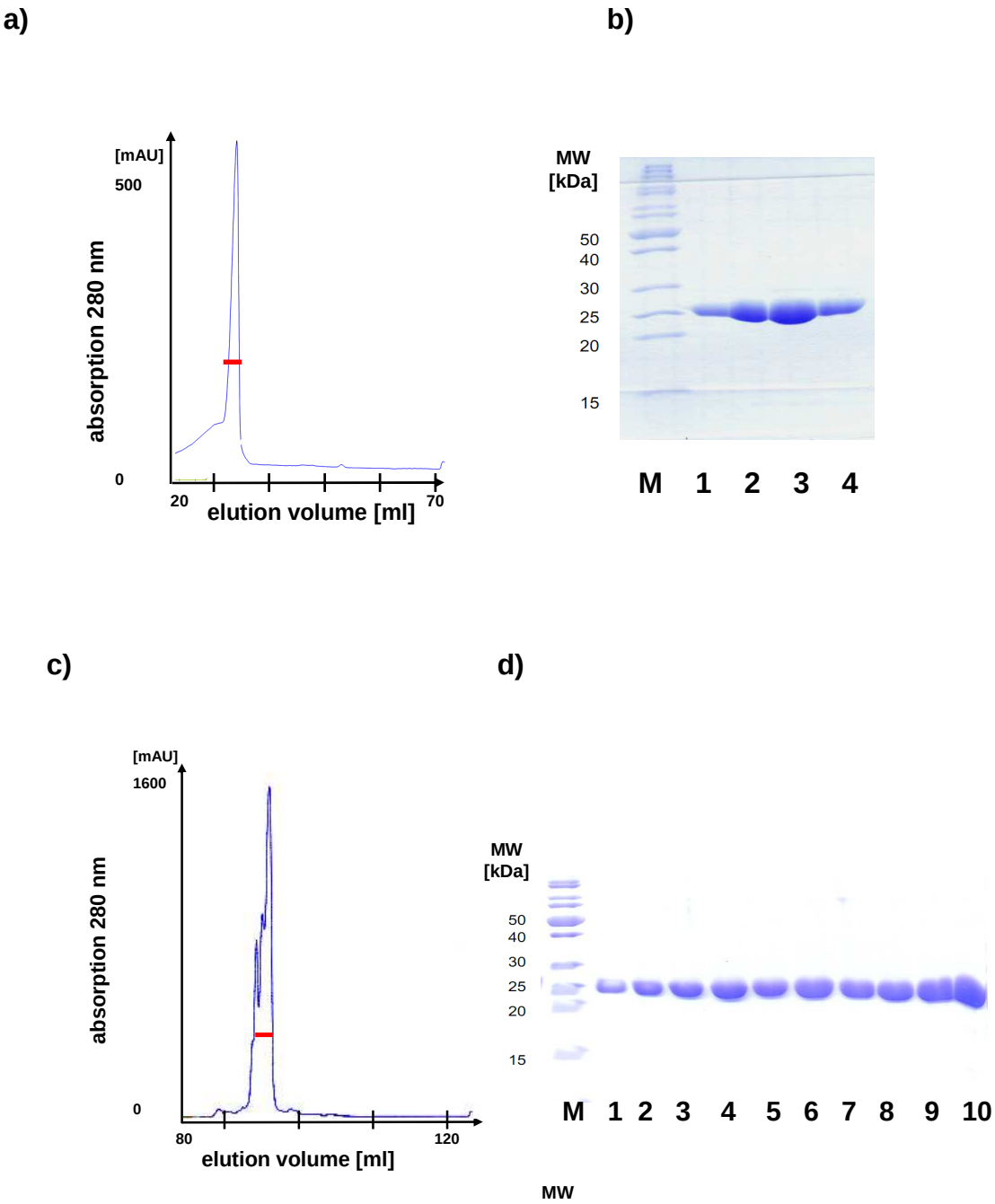


Figure B.10) Ion exchange chromatography and SDS-PAGE analysis of the different constructs of SGT1. **a)** Profile of the cation-exchange chromatography of TPR-CS$_{1\text{-}257}$ with a gradient of NaCl from 0 to 250 mM; **b**) refering SDS-PAGE analysis of the peak fractions as markerked in a); M= molecular weight marker **c)** Elution profile of cation-exchange chromatography of the TPR-CS$_{14\text{-}235}$; **d)** refering SDS-PAGE analysis of the peak fractions, as marked in c). M= molecular weight marker.

B.II.5 Biophysical characterization

B.II.5.a Size and oligomeric state

The size exclusion chromatogram shows one symmetric peak, with minor shoulders, for each of the constructs TPR-CS$_{14-235}$ and TPR-CS$_{1-257}$. SDS-PAGE analysis revealed one clear band, indicating purity of the protein (Figure B.10).

Static light scattering (SLS) confirmed a monomeric state for all of the constructs (Figures B.11 a-d). The obtained molecular weight was compared with the theoretical molecular weight calculated by the program protparam as enlisted in table 1. The full length protein SGT1 showed a molecular weight of 36.5 kDa +/- 1.5% which is in good agreement with the theoretical molecular weight of 38.6 kDa. The observed values for the molecular weight of the other constructs also agree with the predicted values (Table B.1).

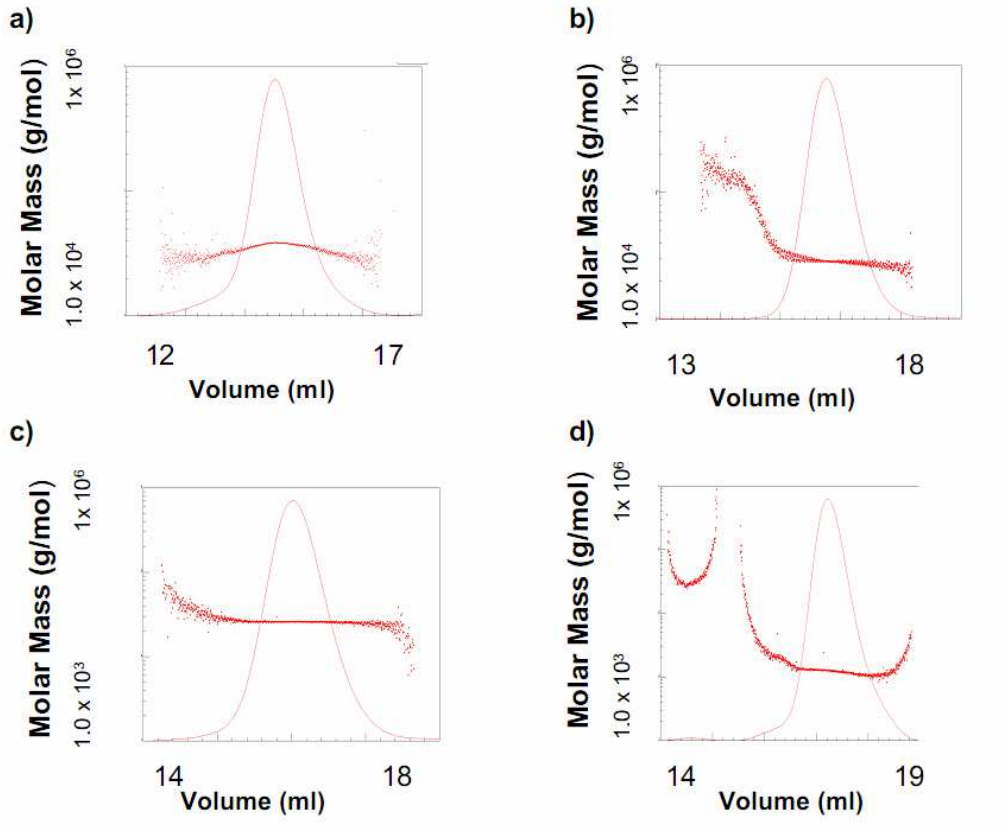


Figure B.11) SLS after SEC The UV280 signal (solid line) and the distribution of the molar mass (dotted line) vs. the elution volume are shown for **a)** the protein SGT1 and the constructs **b)** TPR-CS$_{D1-257}$, **c)** TPR-CS$_{T14-235}$ and **d)** SGS. The main peak of each chromatogram consists of a monomeric species of the respective protein.

Table B.1) Molecular weight of human SGT1 and the constructs TPR-CS$_{1-257}$, TPR-CS$_{14-235}$ and SGS. Shown are the molecular weights as determined by SLS after SEC. The theoretical molecular weight was calculated by the program protparam.

	MW measured	MW theoretical
SGT1	36.5 kDa +/- 1.5 %	38.6 kDa
TPR-CSD1-257	29.2 kDa +/- 0.9 %	29.1 kDa
TPR-CST14-235	25.0 kDa +/- 1.4 %	25.3 kDa
SGS	14.5 kDa +/- 3%	12.6 kDa

B.II.5.b Fluorescence Spectroscopy

Fluorescence spectra were recorded for the full length protein SGT1 and the isolated SGS domain in the presence and absence of the denaturating agent guanidinium-hydrochloride (Figure B.12). Comparing the peak of SGT1 under non-denaturing condition to the peak of SGT1 in the denaturant, a shift of 12 nm is observed. This should confirm the folded state of SGT1, since this difference in the emission is typical for tryptophane residues which are in a hydrophobic environment in folded state and solvent exposed under denatured conditions.

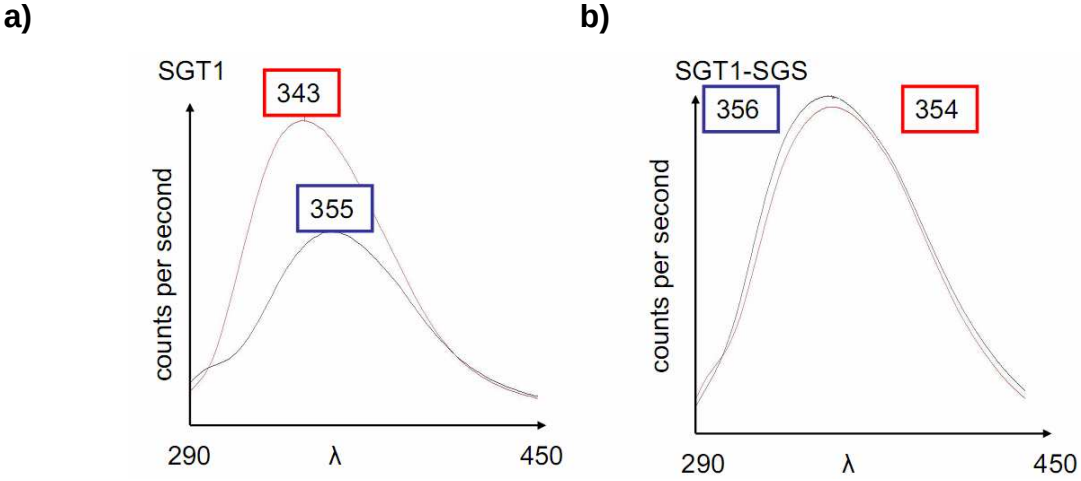


Figure B.12) Fluorescence spectroscopy for **a)** the protein SGT1 and **b)** the SGS domain of SGT1. The images show the emission scans from 290 – 450 nm, after excitation by 280 nm. The measured intensity is given in counts per second. The spectra recorded in non denaturating buffer conditions is presented in blue, the spectra, recorded in denaturating buffer condition is presented in red. The wavelengths of the corresponding peaks are annotated in a box.

The SGS domain on the other hand does not show such a shift. The emission maximum remains at 355 nm, which indicates that the tryptophane resdiues remain solvent exposed.

These results suggest that this domain does not have a well-defined fold in absence of a binding partner.

B.II.5.c Secondary Structure determination

A secondary structure analysis of the different constructs of SGT1 was performed by CD spectroscopy, measuring absorption of circular polarized light in the range of 240 - 200 nm. The amount of the secondary structure was estimated using the three different programs K2D (Andrade, Chacon et al. 1993), K2D2 (Perez-Iratxeta and Andrade-Navarro 2008) and SOM-CD (Unneberg, Merelo et al. 2001).

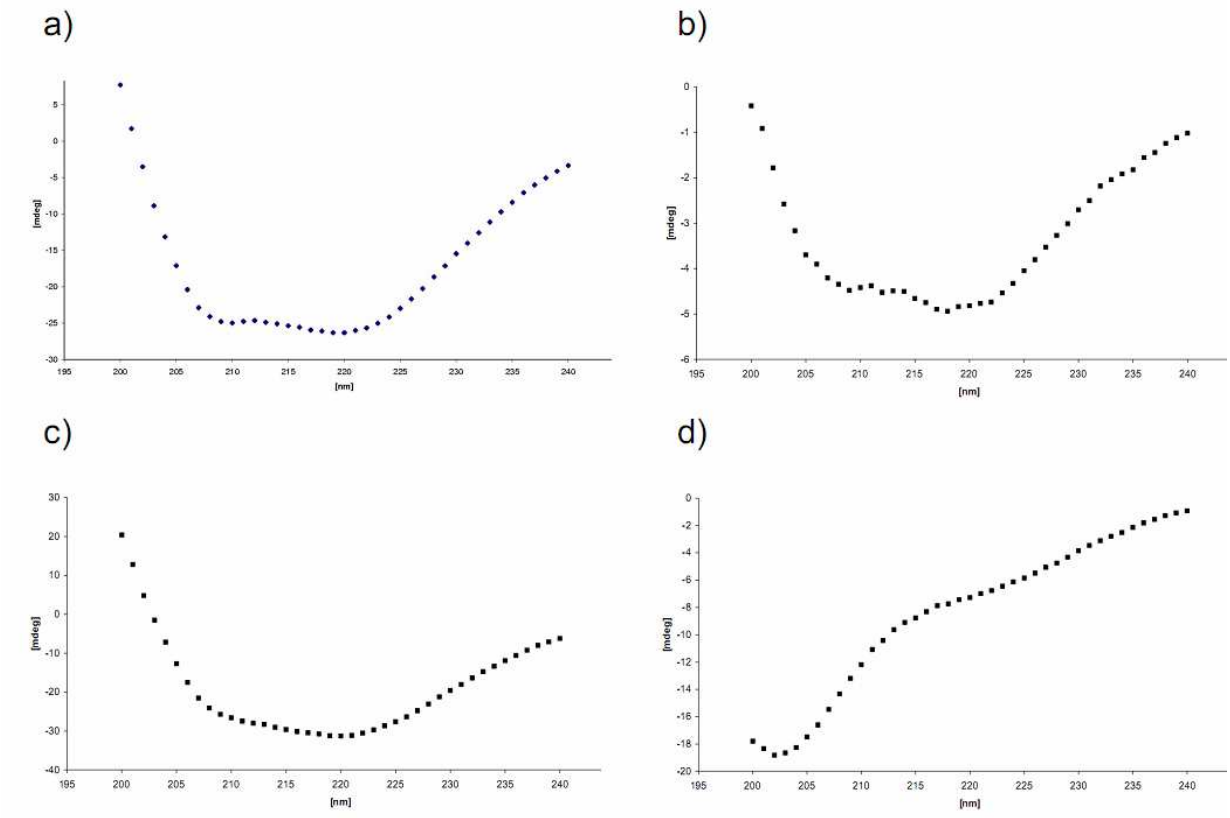


Figure B.13 a-d) Circular Dichroism spectra of **a)** SGT1 **b)** TPR-CSD1-257 **c)** TPR-CST14-235 and **d)** SGS at RT. The absorbance from 200-240 nm is shown, the unit is mdeg. All spectra are measured in 20 mM Phosphate buffer, the reference buffer has already been subtracted. The spectra having minima at 207 nm and 222 nm resemble spectra of alpha-helical proteins.

These three algorithms estimate totally different secondary structure (data not shown). Therefore the spectra were compared to the already published spectra of known secondary structure content.

The CD spectra of the constructs SGT1, TPR-CS$_{1-257}$ and TPR-CS$_{14-235}$ (Figures B.13 a-c) have local minima close to 208 nm and 222 nm, as it is observed for α-helical proteins, (ref 4). These minima are caused by the TPR-domain, which is known to have high α-helical content and shows a comparable pattern (Nyarko, Mosbahi et al. 2007). The obtained spectra might be influence of the known β-sheet content of the CS-domain (Lee, Jacob et al. 2004) shifting the second minima to a lower wavelength (ref 5).

In contrast, the profile of the SGS domain (Figure B.13d) resembles spectra of mainly unstructured proteins (ref 6) having low ellipticity above 210 nm and negative bands near 195 nm (ref 6). This is in accordance with previous observations. Already Lee and coworkers assumed a low degree of folding of the SGS-domain. They compared the CD spectra for this domain to the spectra of the C-terminal domain of calcyclin-binding protein, known to have a limited degree of α-helical content. Since these spectra showed high similarity, therefore they assumed a limited degree of α-helical content for the SGS domain (Lee, Jacob et al. 2004).

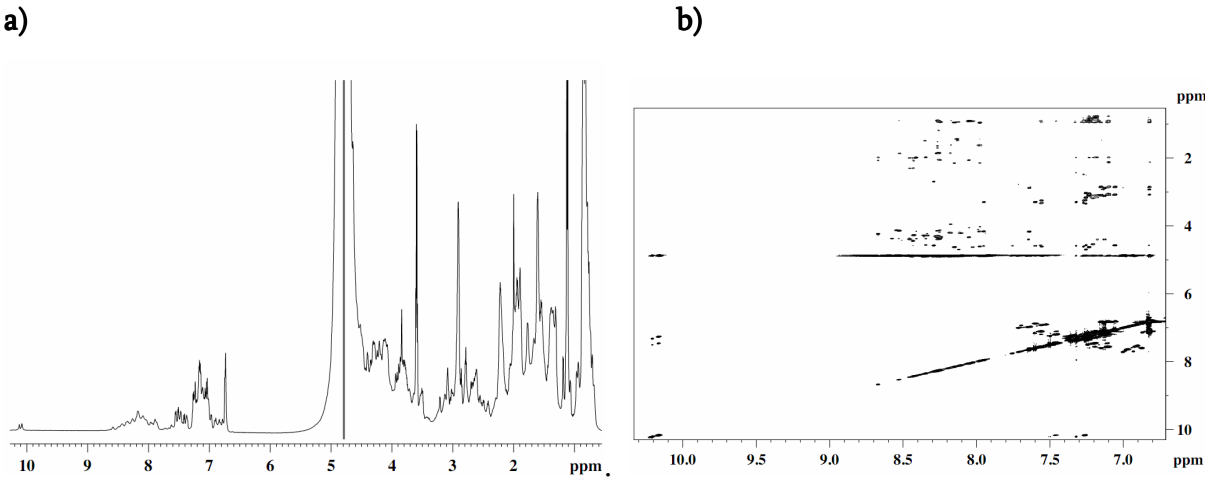


Figure B.14) NMR experiments of the SGS-domain of SGT at 295K a) 1H spectrum of SGS-domain b) NOESY spectrum.

To get an independent estimate for the folding of the SGS-domain, NMR spectroscopy in collaboration with Prof. O. Zerbe (Organic Institute) was performed. The [^{1}H] spectrum of 500 µM SGS shows only little signal dispersion above 8.5 ppm where signals of the

backbone amide protons of folded regions are expected to appear. The signals localized at 10.1-10.2 ppm belong to the side chains of tryptophanes (Figure B.14a). Furthermore, a NOESY spectrum of the SGS domain, which is depicted in Figure B.14b does not show anycross peaks in the chemical shift range of 7.0 - 10.0 ppm. The absenceof these signals indicates the lack of secondary structure elements. Lee and co-workers (2004) concluded from their NMR experiments that this domain does not have a compact well folded globular structure.

B.II.6 Optimization of buffer conditions for crystallization using the Thermofluor method

The homogeneity of a protein sample, as well as its solubility and stability are important factors affecting the success of crystallization experiments.

In order to enhance the possibility of crystallization of SGT1, the initial buffer formulation for the protein was investigated by the Thermofluor method (Ericsson, Hallberg et al. 2006). In this set-up, the protein was thermally denaturated under different buffer conditions. The hydrophobic dye SyproOrange was used to monitor the unfolding of the protein sample. The fluorescence of SyproOrange is quenched in aqueous solution, but the dye becomes fluorescent after binding to hydrophobic patches of proteins. Since protein denaturation makes these hydrophobic patches accessible to the dye, the increase of fluorescence monitors the unfolding of the protein. A shift towards higher melting temperatures (Tm) indicates that a certain buffer stabilizing the conformation of the protein. As seen in table 2 the protein SGT1 was most stable in sodium cacodylate with a Tm of 328 +/-0.2 K). Consequently, this buffer was used as a buffer for protein crystallization.

Table B.2) Thermal denaturation experiments of the protein SGT1 in different buffers covering the pH range from 4.5 - 8. All buffers have a concentration of 100 mM. The different thermal melting points (TM) and the errors are shown.

Buffer (100 mM), pH	Tm [K]	error
Imidazole, 8.0	312.2	0.1
Sodium acetate, 4.5	316.5	0.5
Hepes, 8.0	316.7	0.2
Sodium Phosphate, 7.5	317.4	1.0
Tris, 8.0	317.7	0.2
Sodium citrate, 4.7	318.6	0.1
Sodium acetate 5	319.9	0.0
Tris, 7.5	320.5	0.2
Potassium Phosphate, 5	321.8	0.1
Mes, 5.8	321.9	0.1
Bicine, 8	322.0	0.1
Mes, 6.5	323.1	0.2
Mes 6.2	323.9	0.1
Ammonium acetate, 7.3	324.2	0.2
Hepes, 7	325.2	0.2
Potassium Phosphate, 6.0	325.2	0.1
Sodium Phosphate, 5.5	325.7	0.1
Sodium Phosphate, 6.5	325.8	0.1
Sodium citrate, 5.5	327.8	0.2
Potassium Phosphate, 7	327.8	1.1
Sodium cacodylate, 6.5	328.0	0.2

B.II.7 Crystallization experiments

Up to now, a structure of the full length protein SGT1 is not available. Only the NMR structure of the CS domain of human SGT1 (Lee, Jacob et al. 2004) and the X-ray structure of a complex between the CS domain of SGT1 from *A. thaliana* and the N-terminal domain of HSP90 from barley has been determined (Zhang, Boter et al. 2008). Therefore, a structure of the full length protein would be of great interest. Crystallization experiments of the full length protein SGT1 as well as the constructs TPR-CS$_{1-257}$ and TPR-CS$_{14-235}$ and SGS were performed.

Preliminary ammonium sulphate precipitation screens revealed that SGT1 is most stable in pH conditions ranging from pH 7-8. This was also observed for the other constructs TPR-CS$_{1-257}$, TPR-CS$_{14-235}$ and SGS.

Crystallization was tried at varying protein concentration up to 30 mg/ml, which is the upper concentration limit of the full length protein. The initial buffer formulation was 20 mM Tris, 10 mM NaCl, pH 7.5 or sodium cacodylate pH 6.5. Different crystallization screens, at 4°C and 20°C were exploited. Despite all these efforts, crystals were not obtained.

B.III Discussion

The human protein SGT 1 was purified and characterized using different biophysical methods. Proteolysis resistant fragments, comprising the TPR and CS domain, were defined and corresponding constructs were cloned, purified and characterized.

The full length human protein SGT1 is monomeric and has a certain degree of α-helical content as determined by static light scattering and circular dichroism analysis. This degree could not be determined precisely, neither by our work, nor by previous studies (Lee, Jacob et al. 2004). Interestingly, the oligomerization state depends on the origin of the protein. SGT1 of yeast, barley and Arabidopsis are dimers. Dimer formation is mediated by the TPR domain (Nyarko, Mosbahi et al. 2007).

Since full length SGT1 is highly sensitive to proteolysis, two proteolysis resistant fragments were found by limited proteolysis experiments. These fragments comprised the TPR and the CS domains. They were purified as folded, monodisperse proteins without degradation products and showed defined secondary structure in CD spectroscopy. The CD spectra resemble the recorded spectra for the TPR-domain of barley that is well-known of its high α-helical content (Figure 12) (Nyarko, Mosbahi et al. 2007). The CS domain is mainly composed of β-sheets, which has been determined by NMR (Lee, Jacob et al. 2004) and X-ray crystallography (Zhang, Boter et al. 2008).

Even though we could successfully precipitate transiently expressed NLRP3 by immunoassays (G. Hausammann), using the recombinant SGT1, which requires a functional CS domain, as well as the SGS domain (Mayor, Martinon et al. 2007), our experimental findings confirmed, that the SGS domain without interaction partner is mainly disordered. The CD spectrum of this domain (Figure 12d) resembles the spectrum of a mainly unstructured protein (Greenfield 2006). Lee and (2004) already assumed a limited degree of α-helical content for the SGS domain, by comparing the spectrum with the CD spectrum of the C-terminal domain of calcyclin-binding protein (Nowotny, Spiechowicz et al. 2003). Also, of the NMR spectrum of the SGS domain described here

does not show any noteworthy secondary structure, as it was already indicated by earlier NMR experiments (Lee, Jacob et al. 2004). An induced folding in the presence of a binding partner can be suggested.
Previous analytical ultracentrifugation studies conducted in our lab, showed an unusual elevated frictional ratio (Hausammann 2007). This result and the known proteolysis sensitivity of this protein go along with the lack of secondary structure in the last domain.

It is known that the SGS-domain of yeast SGT1 binds S100A6 (calcyclin) and other members of the S100 protein family (Nowotny, Spiechowicz et al. 2003). In order to investigate the binding characteristics and the possibility of induced folding of the SGS domain, further studies in presence of these interaction partners, are of high interest.

Crystallization experiments were performed with the full length protein and the stable constructs comprising the first two domains, however no crystals could be obtained.
It is known that flexible parts in a protein can impede crystal formation (Wernimont and Edwards 2009). Even though the unstructured domain is excluded in the new constructs, the first two domains are connected by a presumed flexible linker (Lee, Jacob et al. 2004). Co-crystallization of the first two domains with an interaction partner such as DARPins and scFabs holding the domains in a stable conformation would be a promising approach to determine the structure of the first two domains of SGT1 together.

The protein and its constructs serve as valuable tools for further immuno-precipitation studies such as the determination and qualitative characterization of components of the inflammasome and other interaction partners found in the plant pathogen response and in the cell cycle regulation. This has already been realized by immunoprecipitating NLP3 from transiently infected HEK293T cells (Hausammann 2007) and inflammasome components (Mayor, Martinon et al. 2007).

B.IV Material and Methods

B.IV.1 Constructs of human SGT1b

The gene of human SGT1b (NM_006704) and its SGS domain were encoded on the plasmids pGEX6P-1-SGT1b and pGEX6P-1-SGS as a fusion protein to glutathione-S-transferase of *E. coli*. These vectors were a kind gift of J. Tschopp, University of Lausanne.

For the constructs TPR-CS$_{D1-257}$ and TPR-CS$_{T14-235}$, originating from limited proteolysis experiments, PCR was done on the vector pGEX-6-P1-SGT1p with the synthetic oligos (Microsynth, Balgach) Fw & Rev (see table B.3). PCR fragments were subcloned in the vector pGEX-6-P1 (Amersham) using the restriction sites EcoRI/XhoI sites. Sequences were verified by sequencing at Synergene (Schlieren).

.

Table B.3) Primers for cloning the constructs of SGT1.

Construct	primer	Sequence	Tm
TPR-CS$_{D1-257}$	Fw	TATAGAATTCATGGCGGCGGCTGC	64.6
	rev	GAATCTCGAGCTAATAAGGAGATGATGATGGATATAG	63.3
TPR-CS$_{T14-235}$	Fw	ATTGAATTCTTTTTCCAGAGCTTCTCGGATGC	66.5
	Rev	CATCTCGAGCTATTGTTTTGGCGTAGGCAC	66.7

B.IV.2 Overexpression of human SGT1 and the constructs

The proteins were over-expressed in *E.coli* BL21(DE3) cells in 500 ml LB medium supplemented with ampicillin (0.1g/l). Protein production was induced by addition of 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at an OD$_{600}$ of 0.6. The constructs GST-SGT1, GST-TPR-CS$_{D1-257}$ and GST-TPR-CS$_{T14-235}$ were over-expressed at 20°C for 16 h, the construct GST-SGS was over-expressed at 37°C for 4 h. Cells were harvested by centrifugation and stored by -20°C.

B.IV.3 Purification of the proteins

The cell pellets of 4 l bacterial culture were resuspended and lysed in 100 ml Lysis Buffer (10 mM Tris pH 8, 150 mM NaCl, 1 mM DTT, Complete Protease Inhibitor (Roche), 2.5mg/ml DNase, 2mg/ml lysozyme) using the Emulsiflex (Avestin). After the removal of the residual cell debris by centrifugation and filtration with a 0.2 µm filter (Filtropur, Sarstedt) the protein was purified over a GSTrap colum (GE healthcare), washed with Wash Buffer (TBS pH 8, 1 mM DTT) and eluted with Elution Buffer (50 mM Tris, 150 mM NaCl, pH8, 10 mM glutathione, 1 mM DTT). The proteins were dialysed in cleavage buffer (50 mM Tris pH 7, 150 mM NaCl, 1 mM EDTA, 1 mM DTT) and incubated with PreScission protease overnight at 4°C (8 µl stock solution/mg of GST-SGT1). The PreScission-protease, GST-tag and endogenous GST were removed by a second passage over the GSTrap column. After dialysis (membrane Spectra/Por, spectrumlabs, MWCO 12000-14000 Da) into ion-exchange buffer A, ion exchange chromatography was performed. Protein was eluted by a gradient of 50 X column volumes to 50% of buffer B. For the proteins SGT1, TPR-CS$_{D1-257}$ and TPR-CS$_{T14-235}$ cation exchange chromatography was performed with the means of a 1 ml ResS column (GE healthcare). For SGS anion exchange chromatography was performed with the means of a 1 ml ResQ column (GE healthcare). The Anion exchange Buffers were A: 20 mM HEPES pH 7.5, TCEP; B: 20 mM HEPES pH 7.5, TCEP, 250 mM NaCl. Cation exchange buffers were A: 20 mM NaAc pH 5.5, TCEP; B: 20 mM NaAc pH 5.5, TCEP, 250 mM NaCl. Elution was at 6.4 mS/cm for SGT1, 6.31 mS/cm for SGS, 8.3 mS/cm for TPR-CS$_{D1-257}$ and 4.5 mS/cm for TPR-CS$_{T14-235}$. The peak fractions were pooled and concentrated using Amicon Ultra-4 (Millipore) concentrators with a molecular weight cut off (MWCO) of 10 kDa.

B.IV.4 Determination of proteolysis resistant fragments by limited proteolysis

Purified protein SGT1b was kept for 10 days at RT. Aliquots were taken after 0, 4 and 10 days and denaturated in SDS-PAGE buffer. The samples were loaded on a 12 % SDS-PAGE for further analysis or given directly for mass spectrometry analysis.

Purified protein SGT1 was incubated with trypsin in a molar ratio of 1/1000. 5 reactions were started. The reactions were stopped by denaturating in SDS-PAGE loading buffer after 0, 20, 40 and 90 min. The samples were loaded on a 12% SDS-PAGE, before Western Blotting. Bands were cut out for N-terminal sequencing.

Mass spectrometry requires protein samples free of loading buffer. For this purpose the reaction was stopped by TCA precipitation.

One volume of 100% TCA stock was added to 4 volumes of protein sample and incubated at 4°C for 10 min. Then the sample was centrifuged for 5 min at 16 000 g. After removal of the supernatant, the pellet was washed two times with 200 µl of ice-cold acetone. The pellet was finally centrifuged down and dried at 95°C. Mass spectrometry was done at the Protein Analysis Unit of the Functional Genomics Centre Zurich. The sample for TPR-$CS_{T14-235}$ was analyzed by MALDI-MS and Edman-sequencing; whereas the sample for TPR-CS_{D1-257} was analyzed by ESI-MS. Corresponding fragments were determined with the software GPMAW (Peri, Steen et al. 2001).

B.IV.5 Size and oligomeric state determination

Multi angle light scattering was performed after size exclusion chromatography in TBS with the help of a S200 column (GE Healthcare). The Tri-Angle Light Scattering Detector miniDAWN (Wyatt) measured the intensity of the scattered light at the angles 45°, 90° and 135° and the OPTILAB REX rEX was used as a concentration detector, which measured the optical phase difference at the fixed wavelength of 690 nm. An internal calibration run with 100 µl of 2.5mg/ml of BSA was used to detect the correct setting of the dead volume and the delay of the system. The runs with the proteins were done with 100-200 µl of protein at a concentration of 2.0 mg/ml. Curves were evaluated using the ASTRA software.

B.IV.6 Fluorescence Spectroscopy

Fluorescence spectra of 20mg/ml protein sample were recorded in 20 mM phosphate buffer pH 7 in presence and absence of 4 M guanidinium hydrochloride. The sample was excited using light at 280 nm; the emission spectrum was recorded in the range of 290 -

450 nm in order to detect tryptophane fluorescence with the Fluoromax-4 (HoribaJobin Yvon). Reference spectra without protein were subtracted from the measured spectra.

B.IV.7 Circular Dichroism

Circular Dichroism spectra measurements were performed with a Jasco J-715 instrument (Jasco, Japan) at 4 °C. CD spectra were recorded in the wavelength range of 190-245 nm. The cuvette had a pathway length of 0.1 cm and held a volume of 200 µl. The protein was dialyzed overnight in 20 mM phosphate buffer purged under bubbling nitrogen. Measurements were performed at protein concentrations of 20 µM. The reference spectra were carried out with the buffer alone. The sensitivity was 100 mdeg and the spectra were recorded at a speed of 20 nm/min. Three accumulations were acquired for one spectrum.

B.IV.8 NMR

NMR measurements were done by Oliver Zerbe at the Institute of Organic Chemistry of the University of Zurich. The protein sample was measured in 20 mM phosphate, 50 mM NaCl buffer. The protein concentration was 500 µM. 1H spectrum and NOESY spectra were recorded at 295 K.

B.IV.9 Thermal shift assays

7.5 µl of Sypro Orange (300x), 10 µl of SGT1 (60 µM) in 20 mM Tris, 20 mM NaCl and 7.5 µl of testing buffer (Table B.4) were added to the wells of a 96 well PCR plate (BioRAD). The plates were sealed with Optical-Quality Sealing Tape (BioRAD) and heated in an iCycleriQ Real Time PCR Detection System (BioRAD) from 20 to 70 °C in increments of 0.5 °C. The samples were excitated at 490 nm, the fluorescence changes in the wells of the plate were monitored at 575 nm.

The Boltzmann model with the formula (1) was used in order to fit the fluorescence curves obtained with the software Origin (OriginLab). Data points after the fluorescence intensity maximum were excluded from fitting.

(1) $I=[A+(B-A)/(1+e^{(Tm-T)/C}]$

I= fluorescence intensity
T= temperature
A & B = fluorescence intensities before and after the transition, respectively
C= slope factor

Table B.4) Testing buffers were used at a concentration of 100mM:

Buffer, pH	pH
Na acetate	4.5
Na citrate	4.7
Na acetate	5.0
K Pi	5.0
Na Pi	5.5
Na citrate	5.5
Mes	5.8
K Pi	6.0
Mes	6.2
Na Pi	6.5
Na cacodylate	6.5

Buffer, pH	pH
Mes	6.5
K Pi	7.0
Hepes	7.0
NH_4 acetate	7.3
Na Pi	7.5
Tris	7.5
Imidazole	8.0
Hepes	8.0
Tris	8.0
Bicine	8

B.IV.10 Crystallization experiments

Crystallization screens for all of the four constructs were set up at a multitude of conditions, with varying protein concentrations up to the solubility limit, which is around 30 mg/ml in 20mM Tris pH 7.5, 20 mM NaCl. The screening was carried out at the NCCR crystallization facility in small scale in drop sizes of 100-200 nl in 96-well crystallization plates (greiner, crystal quick). All screens were done by vapour-diffusion in sitting drops against different reservoir buffer conditions. The ratio protein/buffer was used at 2:1, 1:1 and 1:2 at incubation temperatures at 4°C and at 20°C.. The screens used were the Clear Strategy (Molecular Dimensions) screen with the buffer range of pH 5.5 to pH 8.5, Sigma-Fluka Factorial Screen, PEG4000/8000, PEG4000/6000, JCSG+ Suite (Nextal), Pact Suite

Crystal (Nextal), Structure Screen I&II (Molecular Dimensions) and the Sigma-Fluka screens #73513 #86684 and #75403.

Pictures of the crystallization trials were taken by the crystal farm immediately after setup and after 1, 4, 7, 10, 13, 28, 58 and 88 days.

B.IV Literature:

Aganna, E., F. Martinon, et al. (2002). "Association of mutations in the NALP3/CIAS1/PYPAF1 gene with a broad phenotype including recurrent fever, cold sensitivity, sensorineural deafness, and AA amyloidosis." Arthritis Rheum **46**(9): 2445-52.

Agostini, L., F. Martinon, et al. (2004). "NALP3 forms an IL-1beta-processing inflammasome with increased activity in Muckle-Wells autoinflammatory disorder." Immunity **20**(3): 319-25.

Alheim, K. and T. Bartfai (1998). "The interleukin-1 system: receptors, ligands, and ICE in the brain and their involvement in the fever response." Ann N Y Acad Sci **840**: 51-8.

Ali, M. M., S. M. Roe, et al. (2006). "Crystal structure of an Hsp90-nucleotide-p23/Sba1 closed chaperone complex." Nature **440**(7087): 1013-7.

Andrade, M. A., P. Chacon, et al. (1993). "Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network." Protein Eng **6**(4): 383-90.

Azevedo, C., S. Betsuyaku, et al. (2006). "Role of SGT1 in resistance protein accumulation in plant immunity." Embo J **25**(9): 2007-16.

Azevedo, C., A. Sadanandom, et al. (2002). "The RAR1 interactor SGT1, an essential component of R gene-triggered disease resistance." Science **295**(5562): 2073-6.

Azpiroz, M. F., E. Rodriguez, et al. (2001). "The structure, function, and origin of the microcin H47 ATP-binding cassette exporter indicate its relatedness to that of colicin V." Antimicrob Agents Chemother **45**(3): 969-72.

Bieri, S., S. Mauch, et al. (2004). "RAR1 positively controls steady state levels of barley MLA resistance proteins and enables sufficient MLA6 accumulation for effective resistance." Plant Cell **16**(12): 3480-95.

Boter, M., B. Amigues, et al. (2007). "Structural and functional analysis of SGT1 reveals that its interaction with HSP90 is required for the accumulation of Rx, an R protein involved in plant immunity." Plant Cell **19**(11): 3791-804.

Breeden, L. and K. Nasmyth (1987). "Similarity between cell-cycle genes of budding yeast and fission yeast and the Notch gene of Drosophila." Nature **329**(6140): 651-4.

Catlett, M. G. and K. B. Kaplan (2006). "Sgt1p is a unique co-chaperone that acts as a client adaptor to link Hsp90 to Skp1p." J Biol Chem **281**(44): 33739-48.

da Silva Correia, J., Y. Miranda, et al. (2007). "SGT1 is essential for Nod1 activation." Proc Natl Acad Sci U S A **104**(16): 6764-9.

Dode, C., N. Le Du, et al. (2002). "New mutations of CIAS1 that are responsible for Muckle-Wells syndrome and familial cold urticaria: a novel mutation underlies both syndromes." Am J Hum Genet **70**(6): 1498-506.

Dubacq, C., R. Guerois, et al. (2002). "Sgt1p contributes to cyclic AMP pathway activity and physically interacts with the adenylyl cyclase Cyr1p/Cdc35p in budding yeast." Eukaryot Cell **1**(4): 568-82.

Ericsson, U. B., B. M. Hallberg, et al. (2006). "Thermofluor-based high-throughput stability optimization of proteins for structural studies." Anal Biochem **357**(2): 289-98.

Fritz, J. H., R. L. Ferrero, et al. (2006). "Nod-like proteins in immunity, inflammation and disease." Nat Immunol **7**(12): 1250-7.

Greenfield, N. J. (2006). "Using circular dichroism spectra to estimate protein secondary structure." Nat Protoc **1**(6): 2876-90.

Hausammann, G. (2007). Expression and Structural Analysis of the Human SGT1. Insitute of Biochemistry. Zurich, University of Zurich. **Master:** 47.

Jones, J. D. and J. L. Dangl (2006). "The plant immune system." Nature **444**(7117): 323-9.

Kadota, Y., B. Amigues, et al. (2008). "Structural and functional analysis of SGT1-HSP90 core complex required for innate immunity in plants." EMBO Rep **9**(12): 1209-15.

Kitagawa, K., D. Skowyra, et al. (1999). "SGT1 encodes an essential component of the yeast kinetochore assembly pathway and a novel subunit of the SCF ubiquitin ligase complex." Mol Cell **4**(1): 21-33.

Koonin, E. V. and L. Aravind (2000). "The NACHT family - a new group of predicted NTPases implicated in apoptosis and MHC transcription activation." Trends Biochem Sci **25**(5): 223-4.

Krippner-Heidenreich, A., R. V. Talanian, et al. (2001). "Targeting of the transcription factor Max during apoptosis: phosphorylation-regulated cleavage by caspase-5 at an unusual glutamic acid residue in position P1." Biochem J **358**(Pt 3): 705-15.

Lee, Y. T., J. Jacob, et al. (2004). "Human Sgt1 binds HSP90 through the CHORD-Sgt1 domain and not the tetratricopeptide repeat domain." J Biol Chem **279**(16): 16511-7.

Martinon, F., L. Agostini, et al. (2004). "Identification of bacterial muramyl dipeptide as activator of the NALP3/cryopyrin inflammasome." Curr Biol **14**(21): 1929-34.

Martinon, F., K. Burns, et al. (2002). "The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-beta." Mol Cell **10**(2): 417-26.

Martinon, F., O. Gaide, et al. (2007). "NALP inflammasomes: a central role in innate immunity." Semin Immunopathol **29**(3): 213-29.

Martinon, F., V. Petrilli, et al. (2006). "Gout-associated uric acid crystals activate the NALP3 inflammasome." Nature **440**(7081): 237-41.

Martinon, F. and J. Tschopp (2004). "Inflammatory caspases: linking an intracellular innate immune system to autoinflammatory diseases." Cell **117**(5): 561-74.

Martinon, F. and J. Tschopp (2005). "NLRs join TLRs as innate sensors of pathogens." Trends Immunol **26**(8): 447-54.

Martinon, F. and J. Tschopp (2007). "Inflammatory caspases and inflammasomes: master switches of inflammation." Cell Death Differ **14**(1): 10-22.

Mayor, A., F. Martinon, et al. (2007). "A crucial function of SGT1 and HSP90 in inflammasome activity links mammalian and plant innate immune responses." Nat Immunol **8**(5): 497-503.

Meylan, E. and J. Tschopp (2005). "The RIP kinases: crucial integrators of cellular stress." Trends Biochem Sci **30**(3): 151-9.

Nadiri, A., M. K. Wolinski, et al. (2006). "The inflammatory caspases: key players in the host response to pathogenic invasion and sepsis." J Immunol **177**(7): 4239-45.

Nowotny, M., M. Spiechowicz, et al. (2003). "Calcium-regulated interaction of Sgt1 with S100A6 (calcyclin) and other S100 proteins." J Biol Chem **278**(29): 26923-8.

Nyarko, A., K. Mosbahi, et al. (2007). "TPR-Mediated self-association of plant SGT1." Biochemistry **46**(40): 11331-41.

Poyet, J. L., S. M. Srinivasula, et al. (2001). "Identification of Ipaf, a human caspase-1-activating protein related to Apaf-1." J Biol Chem **276**(30): 28309-13.

Schwartz, S., Jr., H. Yamamoto, et al. (1999). "Frameshift mutations at mononucleotide repeats in caspase-5 and other target genes in endometrial and gastrointestinal cancer of the microsatellite mutator phenotype." Cancer Res **59**(12): 2995-3002.

Shirasu, K. (2008). "The HSP90-SGT1 Chaperone Complex for NLR Immune Sensors." Annu Rev Plant Biol.

Shirasu, K., T. Lahaye, et al. (1999). "A novel class of eukaryotic zinc-binding proteins is required for disease resistance signaling in barley and development in C. elegans." Cell **99**(4): 355-66.

Shirasu, K. and P. Schulze-Lefert (2003). "Complex formation, promiscuity and multi-functionality: protein interactions in disease-resistance pathways." Trends Plant Sci **8**(6): 252-8.

Unneberg, P., J. J. Merelo, et al. (2001). "SOMCD: method for evaluating protein secondary structure from UV circular dichroism spectra." Proteins **42**(4): 460-70.

Wernimont, A. and A. Edwards (2009). "In situ proteolysis to generate crystals for structure determination: an update." PLoS One **4**(4): e5094.

Ye, Z. and J. P. Ting (2008). "NLR, the nucleotide-binding domain leucine-rich repeat containing gene family." Curr Opin Immunol **20**(1): 3-9.

Zhang, M., M. Boter, et al. (2008). "Structural and functional coupling of Hsp90- and Sgt1-centred multi-protein complexes." Embo J **27**(20): 2789-98.

Chapter - C. Structural Determinants for improved thermal stability of Designed Ankyrin Repeat molecules

C.I) Introduction

Ankyrin repeat motifs can be exploited as stable binding partners for proteins for a wide variety of purposes, such as medical applications, biochemical methods and as partners for co-crystallization for X-ray structure determination. In order to obtain high quality binding partners for proteins of interest Binz and co-workers created Designed Ankyrin Repeat Proteins (DARPins). These proteins are based on a consensus design, and are characterized to be extremely stable (Binz, Stumpp et al. 2003).

Experimental studies and Molecular Dynamics simulation have demonstrated that the carboxy-terminal capping repeat (C-cap) of these DARPins is denaturing first, followed by the ankyrin body (Yu, Kohl et al. 2006; Interlandi, Wetzel et al. 2008). Several mutations have been proposed by molecular dynamics (MD) simulations in order to fortify this C-cap, thereby increasing the overall stability of the protein. The stability of the mutants carrying the proposed mutations were determined by thermal and chemical unfolding kinetics of the NI_1C, a designed full consensus ankyrin repeat protein, possessing only one internal repeat (Interlandi, Wetzel et al. 2008).

Also for the use of these binding proteins as a crystallographical tool, it is of outstanding interest to fortify the C-cap. It has been observed, that the electron density of the C-cap is poorly defined in most of the known DARPin structures. For the ankyrin F8, a selected DARPin inhibitor of caspase 2, no electron density has been found for residues 154 to 166, neither in the complex with caspase-2 nor crystallized without a binding partner (Schweizer, Roschitzki-Voser et al. 2007). Better crystallographic results were achieved by using a DARPin with a deleted C-cap for crystallizing the kinase domaine of the Polo-like kinase1 (Plk1). Here the absence of the C-cap enhanced the binding of the DARPin to the Plk1 by the formation of a closed tail to tail contact between the two DARPins bound to the protein. (Bandeiras, Hillig et al. 2008).

The focus of this project in my PhD thesis was on the following questions: What are the structural determinants for the improved stabilities of the C-cap mutants. The crystal structures of the DARPin C-cap mutants NI_1C mutant 4, NI_3C mutant 5, and NI_3C mutant 6 were solved by X-ray crystallography.

The crystal structures were the basis for a detailed analysis of structural factors contributing to the higher stability. The important characteristics of the interfaces of the ankyrin body to the C-cap, such as surface area, shape complementarity, hydrogen-bonding network and salt bridges were investigated in detail.

The mutants 5 and 6 possess the highest number of point mutations and reveal conformational changes in the positioning of the C-cap, which are described as a rigid body movement of the C-cap towards the body of the ankyrin structure. Consequently the number of hydrophobic contacts and H-bonds are increased and fortify the interactions in the interface between the ankyrin body and the C-cap (NIII-C). The geometrical shape and the buried surface area of the interrepeat interfaces is altered, a better stacking of the C-cap to the last internal repeat is enhanced.

Title: **Structural Determinants for improved thermal stability of Designed Ankyrin Repeat molecules.**

Authors: Michaela A. Kramer, Svava K. Wetzel, Andreas Plückthun, Peer R.E. Mittl & Markus G. Grütter*

Address: Department of Biochemistry, University of Zürich, Winterthurerstr. 190, 8057 Zürich, Switzerland

Running title: Structures of DARPin C-cap mutants

Text: 19 pages, 2 figures, 2 tables

Abbreviations: DARPin, Designed Ankyrin Repeat Protein

GdnHCl, guanidinium hydrochloride

rmsd, root mean square deviation

VdW, Van der Waals

NCS, none-crystallographic symmetry

ASU, asymmetric unit

Abstract

Designed ankyrin-repeat proteins (DARPins) serve as very versatile protein recognition modules, because DARPins against almost any target molecule can be obtained by using library based combinatorial techniques. Although DARPins are already extremely stable molecules, heat denaturation experiments, structural studies and molecular dynamics simulations suggested that the unfolding of the C-terminal capping repeat limits the stability of the initial DARPin design. Using the consensus sequence of the internal repeats as a guide, several point mutations were suggested to optimize the design of the C-terminal capping repeat. These mutations were introduced and the crystal structures of the more stable DARPins NI_1C mutant 4, NI_3C mutant 5 and NI_3C mutant 6 were refined at 2.2 Å, 2.1 Å and 1.8 Å resolution, respectively. In the optimized design the C-terminal capping repeat has moved as a rigid body towards the preceding internal repeat. This movement generates an increased buried surface area and superior surface complementarities, which explains the improved thermal stability compared to the initial scaffold design.

Introduction

Repeat proteins are the most abundant class of naturally occurring protein-protein interaction modules. They are found in all kingdoms of life and constitute rather simple architectures, consisting of structurally identical repeats that are stacked together to form an elongated none-globular protein domain. Multi-repeat proteins serve as unique templates for protein engineering experiments, because their stabilities are dominated by

intra-repeat interactions and interactions between neighbouring repeats. (Bork 1993; Andrade, Perez-Iratxeta et al. 2001)

Several classes of multi repeat proteins exist. Besides armadillo-, tetratricopeptide- and leucine-rich repeat proteins (Main, Xiong et al. 2003; Stumpp, Forrer et al. 2003; Parmeggiani, Pellarin et al. 2008) members from the family of ankyrin repeats (AR) have probably been most often used as protein engineering templates. The AR consists of two anti-parallel α-helices and a short β-turn. Two different approaches have been applied to design AR proteins by protein engineering. (Mosavi, Minor et al. 2002; Binz, Stumpp et al. 2003) Both approaches relied on the design of consensus sequences that were assembled based on large collections of naturally occurring AR sequences. Although both approaches resulted in stable AR proteins with different numbers of repeats, the design of Mosavi et al. (2002) yielded proteins that were only soluble at the acidic pH range, whereas the approach of Binz et al. (2003) yielded proteins that were soluble over a larger pH range. The incorporation of N- and C-terminal capping repeats greatly improved the solubility of these DARPins and they were subsequently used as very versatile and selective protein-protein interaction modules for biochemical and biophysical studies. Crystal structures of isolated DARPins and DARPins in complex with various target proteins have been determined. (Kohl, Binz et al. 2003; Amstutz, Binz et al. 2005; Kohl, Amstutz et al. 2005)

DARPins that were generated following the design of Binz et al. (2003) are described as NI_xC molecules, where *N*, *I*, and *C* represent the N-terminal, internal and C-terminal repeats and *x* indicates the number of internal repeats. Internal repeats consist of 33

residues each; 27 of them serve as framework residues, which are conserved in all internal repeats, and 6 residues are selected by combinatorial techniques, such as ribosome display, to recognize the target molecule. In full-consensus DARPins the amino acid sequences of all internal repeats are exactly identical.

DARPins are extremely stable molecules. Generally the stabilities of full-consensus DARPins increase with the number of internal repeats and DARPins with more than two internal repeats can not be denatured by heat treatment alone. (Interlandi, Wetzel et al. 2008; Wetzel, Settanni et al. 2008) Depending on the ionic strength the full-consensus NI_3C possesses melting temperatures between 74 °C and 80 °C even in the presence of 4 M GdnHCl. (Merz, Wetzel et al. 2008; Wetzel, Settanni et al. 2008) Comparison of several DARPin crystal structures revealed that the C-terminal capping repeats possess the largest structural flexibilities. (Kohl, Binz et al. 2003; Binz, Kohl et al. 2006; Merz, Wetzel et al. 2008) Molecular dynamics calculations of the none full-consensus E3_5 and E3_19 (Yu, Kohl et al. 2006) and several full-consensus NI_1C and NI_3C DARPin structures (Interlandi, Wetzel et al. 2008) showed, that the C-terminal capping repeat possesses the lowest structural stability. Indeed, NI_3C showed a bi-phasic behaviour in GdnHCl induced unfolding experiments, whereas NI_3 showed a clear mono-phasic behaviour. Since NI_3 lacks the C-terminal capping repeat, the pre-transition in NI_3C unfolding experiments was interpreted as the premature unfolding of the C-terminal capping repeat. (Interlandi, Wetzel et al. 2008)

Since the C-terminal capping repeats prevent the formation of insoluble aggregates and in order to improve the temperature stabilities of DARPin molecules even further, Interlandi

et al. (2008) suggested six mutations to improve the packing between the C-terminal capping repeat and the preceding internal repeat (Table 1). In contrast to the design of Binz et al. (2003), where the C-terminal capping repeat was taken from the guanine-adenine binding protein β1, the new design was inspired by the consensus sequence of the internal repeats. To investigate the structural consequences of those mutations we determined the crystal structures of NI_3C mutants 5 and 6 as well as NI_1C mutant 4 and compared them to the crystal structure of the ancestor molecule NI_3C. (Merz, Wetzel et al. 2008) Interestingly many full-consensus DARPins are prone to form twinned crystals, which was overcome by using ammonium sulphate as the crystallization agent.

Results

Structure of NI_1C mutant 4

The crystal structure of NI_1C was determined at 2.2 Å resolution and folds into a compact globular molecule with overall dimensions of 17 · 24 · 25 Å (Fig. 1a). The structure reveals clear electron density for amino acids 13 to 103, which are sub-divided into the N-terminal capping repeat (residues 13-43), the internal repeat (residues 44-76) and C-terminal capping repeat (residues 77-103). Although several crystal structures of DARPins with three internal repeats have been determined, the present structure of NI_1C mutant 4 serves as the prototype structure for a DARPin with just one internal repeat. The NI_1- (residues 13 to 76) and I_1C repeat pairs (residues 43 to 100) of NI_1C mutant 4 superimpose perfectly onto the corresponding N- and C-terminal pairs of NI_3C wild type with rmsds of 0.47 Å (66 Cα-atoms) and 0.46 Å (54 Cα-atoms), respectively. The N-, I- and C-terminal repeats of NI_1C reveal average B-factors of 19.1 $Å^2$, 24.7 $Å^2$ and 22.6 $Å^2$,

respectively. Thus, in contrast to previously determined structures of DARPins with three internal repeats, there is just a very moderate B-factor gradient present in NI_1C mutant 4, suggesting that the mutations effected the flexibility of the C-terminal capping repeat.

The C-terminal capping repeat of NI_1C mutant 4 deviates from the capping repeat of NI_3C wild type by a three amino acid extension of the second C-terminal capping repeat helix (helix B). As it has been anticipated during the design process, residues Lys_{101}-Ala-Ala_{103} adopt an α-helical conformation and participate in the hydrogen bonding network of helix B. The side chain of Lys101 serves as a C-terminal helix-cap, but forms no direct hydrogen bonds with the C-terminus of helix B. Instead Lys101 forms a surface exposed salt bridge with the side chain of Glu97. Furthermore, the Cα-atom of Ala103 is located at VdW distance to the side chain of Val73 from the preceding internal repeat.

NI_1C mutant 4 shows a markedly different side chain packing at the interface between internal- and C-terminal repeats (Fig. 1b). In NI_1C mutant 4 the side chain of His52 forms a salt bridge with the side chain of Asp77, whereas in NI_3C the equivalent His118 interacts with Thr115. The movement of His52 causes the rotation of the side chains of Ile86, Ser87 and Leu95, since they are all at VdW distance to one another. In NI_1C mutant 4 Ser87-OG forms a hydrogen bond with Ala83-O (3.05 Å). Since none of the rotated side chains are in direct contact with the Lys_{101}-Ala-Ala_{103} extension of helix B, the observed movements could be the consequence of different crystallization conditions. NI_1C mutant 4 was crystallized at pH 5.0 and in the presence of 30 % methanol, whereas NI_3C wild type was crystallized at pH 8.5 and in the absence of methanol. Neither the extension of helix B nor the different side chain conformations of His52, Ile86, Ser87 and

Leu95 have a significant effect on the conformation of the capping repeat main chain. Thus, the spatial orientations of the C-terminal capping repeats relative to the last internal repeats are almost identical in the structures of NI_3C wild type and NI_1C mutant 4.

Structures of NI_3C mutants 5 and -6

NI_3C mutants 5 and -6 crystallized with 4 and 3 molecules in the asymmetric units and were refined at 2.1 Å and 1.8 Å resolution, respectively. The structures of the N-terminal capping repeats and the internal repeats of both mutants were almost identical to the corresponding repeats of the NI_3C wild type (Fig. 2a). Superposition of Cα atoms of residues 13 to 76 from 2QYI (Merz, Wetzel et al. 2008) onto mutants 5 and -6 revealed rmsds of 0.31±0.01 Å and 0.24±0.04 Å, respectively. Larger structural changes are seen at the interfaces between the last internal repeats and the C-terminal capping repeats. Superposition of residues 110 to 166 (corresponding to the last internal repeat and the C-terminal capping repeat) revealed rmsds of 0.78±0.01 Å and 0.68±0.07 Å for mutants 5 and -6, respectively. In mutants 5 and -6 the C-terminal capping repeats have moved as rigid bodies towards the last internal repeats (Fig. 2b). The largest movements are seen for residues 154 to 160, which form the loop between helices A and -B of the C-terminal capping repeat. These residues have moved approximately 2 to 3 Å towards the last internal repeat (Fig. 2b). The redesign also increased the buried surface area of the interface between the last internal repeat and the capping repeat from 625 $Å^2$ in NI_3C wild type to 720±6 $Å^2$ and 723±6 $Å^2$ in mutants 5 and -6, respectively. Simultaneously mutants 5 and -6 show improved surface complementarities compared to the wild type interface (Table 1).

The rigid body movement of the capping repeat is a consequence of the new interface design. Mutant 5 deviates from the NI_3C wild type (Table 1) by a 3 amino acid extension of helix B (similar to NI_1C mutant 4) and the additional point mutations Ala149Pro, Ile152Leu, Ser153Ala, Leu161Ile, and Ile164Val (equivalent to Ala83Pro, Ile86Leu, Ser87Ala, Leu95Ile, and Ile98Val in reference (Interlandi, Wetzel et al. 2008)). Mutations Ala149Pro, Ser153Ala, and Leu161Ile occupy buried hydrophobic pockets at the interface between the last internal repeat and the C-terminal capping repeat, whereas mutations Ile152Leu and Ile164Val are partially solvent accessible (Fig. 2b). The mutation Ala149Pro is located at the N-terminus of helix A and replaces an internal water molecule in the loop region between helix A and the preceding β-strand. The Ser153Ala mutation was inserted because it increases the helix propensity (Interlandi, Wetzel et al. 2008). The crystal structure of mutant 5 shows that this mutation abrogates the H-bond between Ser153-OG and Phe150-O at the interface between helices A and B. However, this mutation seems to have almost no effect on the distances between helices A and B, which are 5.3 Å and 5.4 Å in wild type and mutant 5, respectively (distances between Cα-atoms at positions 153 and 162). In NI_3C mutant 5 the side chains of Leu152, Ile161 and Val164 are penetrating deeper into the corresponding cavities with the consequence that the C-terminal capping repeat is pulled towards the last internal repeat. This reorientation of the capping repeat allows the formation of the hydrogen bonds Asn158-ND2...Ala121-O (2.81±0.05 Å) and Asn156-ND2...Leu152-O (2.88±0.14 Å, in chains A, B, and D) in mutant 5, which are not seen in the wild type structure. In chain C Asn156 participates in the formation of a crystal contact to Gly91-O of chain A.

NI_3C mutant 6 contains the additional mutations Asp155Arg, Asn156Glu, and Asn158His (equivalent to Asp89Arg, Asn90Glu, and Asn92His in reference (Interlandi, Wetzel et al. 2008)) relative to mutant 5. Mutations Asp155Arg and Asn156Glu were inserted, because it was expected that the side chains of the newly introduced residues would form a stabilizing salt bridge as it was seen in the internal repeats of NI_3C wild type. (Interlandi, Wetzel et al. 2008) However, in all chains of NI_3C mutant 6 the distances between the guanidinium groups of Arg155 and the carboxylate groups of Glu156 were larger than 4 Å, suggesting that no salt bridges were formed under the given crystallization conditions. Furthermore, the Asn156Glu mutation abrogates the hydrogen bond between the Asn156 side chain and the main chain of Leu152, which was observed in NI_3C mutant 5. Mutation Asn158His allows the formation of the hydrogen bond His158-ND1...Ala122-O (2.91±009 Å) that links the C-terminal capping repeat to the last internal repeat as it was already seen in mutant 5 for Asn158 (Fig. 2c). In general residues 155 to 158 are rather flexible in the NI_3C mutant 5 and -6 crystal structures and consequently show different conformations and weak electron densities in some NCS related molecules.

Impact of precipitating agent on crystal twinning

Some full-consensus DARPins are prone to form twinned crystal lattices. NI_3C mutant 5 crystallized from sodium citrate/PEG conditions and initial data sets were indexed assuming a hexagonal crystal lattice. Since the structure could not be refined using the hexagonal settings and because data statistics suggested pseudo-merohedral twinning, the structure was refined assuming the trigonal space group $P3_121$, twin law (-h, -k, l) and 4 molecules in the asymmetric unit. Twining is often difficult to assign in the presence of

none-crystallographic symmetry. Therefore we used the RvR plot (Lebedev, Vagin et al. 2006) to investigate the agreement between the observed diffraction data and the refined structure. Values for R^{obs}_{twin} of 0.17 and for R^{calc}_{twin} of 0.30 confirmed our assumption of pseudo-merohedral twinning in the presence of rotational NCS. The relatively low R^{calc}_{twin} value agrees very well with a twinning fraction of 0.374.

Besides NI_3C mutant 5 also the NI_3C wild type (Merz, Wetzel et al. 2008) initially yielded twinned crystals when sodium citrate was used as precipitating agent (PREM unpublished observation). For the wild type NI_3C twinning was overcome by replacing sodium citrate against ammonium sulphate. NI_1C mutant 4 and NI_3C mutant 6 also crystallized from ammonium sulphate solutions and these crystals were not twinned. Analysis of the crystal lattices revealed that sodium citrate causes the formation of crystal contacts involving N-cap/N-cap* and C-cap/C-cap* stacking interactions, whereas all crystals originating from ammonium sulphate conditions only show N-cap/C-cap* stacking interactions (* refers to the symmetry related molecule). In addition well-defined sulphate ions were identified in the electron density maps during the refinement processes. In NI_1C mutant 4 one sulphate ion was observed in a crystal contact close to helix A from the internal repeat (Fig. 1a), whereas in NI_3C mutant 6 three sulphate ions occupy similar positions like in NI_3C wild type (Fig. 2a). However, in NI_3C mutant 6 the sulphate ions are not directly involved in crystal contacts.

Discussion

Redesign of the interface between the last internal repeat and the C-terminal capping repeat significantly improved stability of full-consensus DARPins (Interlandi, Wetzel et

al. 2008; Wetzel, Settanni et al. 2008) and the crystal structures of NI_1C mutant 4, NI_3C mutant 5 and NI_3C mutant 6 reveal the molecular determinants for this stabilization. The mutation Ala149Pro places a proline residue at the N-terminus of helix A. At this position proline participates in the hydrogen bonding network of helix A and improves helix formation due to its restricted ϕ-value. Furthermore mutations Ile152Leu, Leu161Ile, and Ile164Val enable a tighter packing of the C-terminal capping repeat onto the last internal repeat. Altogether these mutations allow for a reorientation of the capping repeat that in a first approximation moves as a rigid body. This movement can be described as a 7° hinge movement of the capping repeat towards the last internal repeat and an equal rotation along the axis of the stacked repeats. The additional hydrogen bond between Asn158-ND2 and Ala121-O is a consequence of this reorientation because the capping repeat adopts the same conformation in mutant 6 although this hydrogen bond is disrupted by the Asn158Glu mutation. Thus, the C-terminal capping repeats of mutants 5 and -6 provide larger interface areas, improved surface complementarities and consequently superior shielding of the hydrophobic core against the solvent compared to NI_3C wild type. The improved stability is reflected by significantly weaker B-factor gradients. The NI_3-parts (residues 13 to 142) and C-terminal repeats (residues 143 to 169) reveal average B-factors of 31.7±0.4 $Å^2$ and 33.1±1.6 $Å^2$ in mutant 5, 31.5±2.9 $Å^2$ and 39.4±3.5 $Å^2$ in mutant 6, but 23.0 $Å^2$ and 34.1 $Å^2$ in NI_3C wild type.

Even though none of the mutations discussed above have been implemented in NI_1C mutant 4, this mutant shows improved structural stability over NI_1C wild type as judged by GdnHCl-induced unfolding experiments (Interlandi, Wetzel et al. 2008) and almost no B-factor gradient. The improved stability of the capping repeat can be addressed to the

extension of helix B, which provides a C-terminal helix cap (Lys101) and additional VdW contacts with the preceding internal repeat. Whether the conformation of His52 also contributes to the improved stability remains questionable, because its movement can't be addressed to the extension of helix B. The altered side chain rotamer of His52 could be a consequence of modified crystallization conditions, particularly the addition of methanol. NI_1C was crystallized under acidic conditions (pH 5.0) and in the presence of 30 % methanol, whereas all NI_3C crystals were obtained in the absence of methanol and under different pH conditions (mutant 5 at pH 4.1, mutant 6 at pH 5.5, wild type at pH 8.5). Since all NI_3C structures reveal similar conformations of His118 (the equivalent to His52) its side chain conformation is independent from the pH of the crystallization condition.

In all full-consensus DARPin structures the N- and C-terminal caps are involved in crystal contacts. In the crystal lattices of none-twinned NI_3C wild type, NI_1C mutant 4 and NI_3C mutant 6 crystal contacts are formed by N-cap/C-cap* stacking interactions, whereas in the crystal lattice of twinned NI_3C mutant 5 crystal contacts are formed by N-cap/N-cap* and C-cap/C-cap* stacking interactions. The crystal lattices of NI_3C wild type and NI_3C mutant 6 are almost isomorphous, since both crystals belong to space group $P6_1$ with a unit cell c-axis of approximately 50 Å and similar super-helical crystal packings of NI_3C molecules. Thus, in the case of full-consensus DARPins the precipitating agent ammonium sulphate causes the formation of crystal lattices involving N-cap/C-cap* stacking interactions and these lattices seem to be less prone to form twinned crystals. Furthermore, an implication of the crystal lattice on the interface between the last internal repeat and the C-terminal capping repeat can be ruled out,

because NI_3C wild type and mutant 6 show significantly different interfaces although the crystal contacts are almost identical, *vice versa* NI_3C mutants 5 and -6 show very similar interfaces although the crystal contacts are significantly different.

The crystal structures of NI_1C mutant 4, NI_3C mutant 5 and NI_3C mutant 6 presented above, illustrate the success of the consensus sequence approach for the design of new C-terminal capping repeats. (Interlandi, Wetzel et al. 2008) From a pure structural point of view the extension of helix B by Lys101, Ala102 and Ala103, the replacement of alanine against Pro83 at the N-terminus of helix A and the introduction of the optimized interface residues Leu86, Ile95 and Val98 contribute significantly to the improved stability of the new C-cap design. Whether the increased helix propensity of the Ser87Ala mutation improves the stability can not be deduced from the crystal structures alone. Ser87 could still be beneficial for the stability of helix A. The mutations Asp89Arg and Asn90Glu were introduced, expecting that the Arg89...Glu90 salt bridge would stabilize the cap. Under the given crystallization conditions for NI_3C mutant 6 no such interaction was observed. Thus, although the new C-terminal capping repeat design of Interlandi et al. (2008) is superior over the initial design there is still room for further improvement.

Material and Methods

Protein purification and crystallization

NI_1C mutant 4 and NI_3C mutants 5 and -6 were expressed in *E. coli* and purified as described previously. (Binz, Amstutz et al. 2004) Size-exclusion chromatography was performed in 10 mM HEPES, pH 6.5, 20 mM NaCl on a Hi-Load 26/60 Superdex-75 column (GE Healthcare) mounted on an Äekta Prime system (GE Healthcare). The

proteins were concentrated to 30 mg/mL using Centricon (Millipore USA) ultra-filtration devises (3 kDa molecular mass cutoff). Crystals were grown using the sitting drop-vapor diffusion method at 20°C in 24-well crystallization plates. Proteins and reservoir solutions were mixed at 2 µL to 1µL ratios. Crystallization conditions are given in Table 2. Crystals were soaked in the reservoir buffers supplemented with 30 % glycerol as cryo-protectant and flash-cooled in a nitrogen stream at 100 K.

X-ray data collection and refinement

X-ray diffraction data were measured at the PX6 beam-line at the Swiss Light Source (Villingen, CH) using the Pilatus detector. For each structure 360 frames were recorded with an oscillation range of 0.5° per frame. Data were processed using the program XDS. (Kabsch 1993) Twinning of X-ray diffraction data was analyzed using the program phenix.xtriage. (Adams, Grosse-Kunstleve et al. 2002) Initial phases were determined by molecular replacement using program Phaser (McCoy, Grosse-Kunstleve et al. 2007) with the structures of the full-consensus NI_3C wild type (pdb code: 2QYJ)(Merz, Wetzel et al. 2008) or the three-repeat ankyrin repeat protein $3CA_{1A2}$ (pdb code: 2ZGD)(Kelly, McDonough et al. 2009) as search models. Structures were refined using the programs PHENIX (Adams, Grosse-Kunstleve et al. 2002) and COOT (Emsley and Cowtan 2004). NI_3C mutant 5 was refined using the twin law (-h, -k, l) with a twinning fraction of 0.374.

The qualities of the structures were evaluated with the programs MOLPROBITY and PROCHECK (Laskowski, Moss et al. 1993; Davis, Leaver-Fay et al. 2007). To investigate the intra-molecular shape complementarities we used the program SC (Lawrence 1993), which was originally developed to calculate the shape

complementarities between independent subunits. Therefore separate coordinate files were prepared by splitting the DARPin structures at the repeat boundaries (residues 76/77 for NI_1C and residues 142/143 for NI_3C). PYMOL (DeLano 2002) was used to prepare figures.

Accession numbers

Coordinates and structure factors have been deposited at the Macromolecular Structure database with accession number XXX (NI_1C mutant4), XXX (NI_3C mutant 5), and XXX (NI_3C mutant 6),

Figure legends

Figure 1. (a) Ribbon diagrams of NI_1C mutant 4. N-terminal capping repeat, internal repeat and C-terminal capping repeat are coloured salmon, green and blue, respectively. The N-terminus and the last amino acid of every repeat is labelled. (b) Wall-eye stereo plot of the superposition of the last internal repeat and the C-terminal capping repeat of NI_1C (light blue) onto the corresponding repeats of NI_3C wild type (magenta, pdb-id: 2qyj). Amino acid numbering refers to NI_1C with the exception of Thr115, which is shown in italics.

Figure 2. (a) Ribbon diagrams of NI_3C mutant 6. Colour coding according to Fig. 1a. (b) Wall-eye stereo plot of the superposition of NI_3C-mutant 5 (cyan) onto NI_3C wild type (magenta). Residues 13-142 of mutant 5 are depicted as a molecular surface (grey) and the C-terminal capping repeats as Cα-traces. Residue nomenclature follows the sequence of NI_3C mutant 5. (c) Superposition of C-terminus of NI_3C mutant 6 (green carbon atoms) onto NI_3C mutant 5 (blue carbon atoms). Different shades of blue indicate the repeat boundaries (light blue: internal repeat; cyan: C-terminal capping repeat).

Tables

Table 1. Sequence and structural parameters of full-consensus DARPins

Repeat	Sequence‡	Tm * (C)	Buried area (Å²)†	Surface complementarities †
N-cap	--DLGKKLLEAARAGQDDEVRILMANGADVNAK			
internal	DKDGYTPLHLAAREGHLEIVEVLLKAGADVNAK			
C-cap				
wild type	DKFGKTAFDISIDNGNEDLAEILQ--------	60°	625	0.705
mutant 4	DKFGKTAFDISIDNGNEDLAEILQKAA-----	64°	654	0.711
mutant 5	DKFGKTPFDLAIDNGNEDIAEVLQKAA-----	77°	720 ± 6	0.801
mutant 6	DKFGKTPFDLAIREGHEDIAEVLQKAA-----	77°	723 ± 6	0.802

‡ Residues participating in α-helices are indicated in bold and mutated residues in red characters.

* Melting temperature of NI_1C wild type and mutants, taken from (Interlandi, Wetzel et al. 2008).

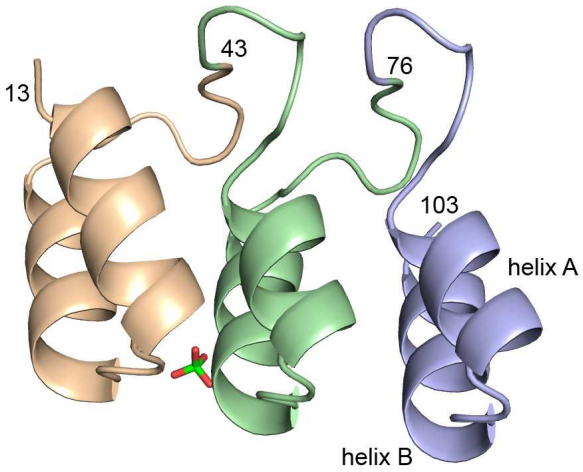
† Buried surface area and surface complementarities between last internal repeat and C-terminal capping repeat. Errors represent standard deviations between values of NCS-related molecules.

Table 2. Crystallization parameters and refinement statistics

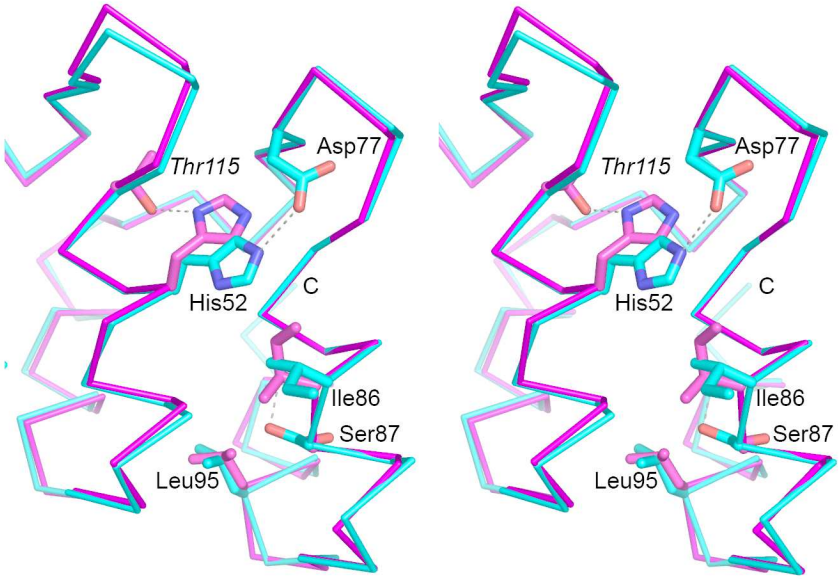
	NI_1C mutant 4	**NI_3C mutant 5**	**NI_3C mutant 6**
Data collection			
space group	C2	$P3_121$	$P6_1$
molecules/asymmetric unit	1	4	3
unit cell parameters (Å)	a=77.15, b= 35.33, c=32.40 α=90°, β=114.43°, γ=90°	a=62.05, b=62.05, c=270.18 α=90°, β=90°, γ=120°	a=128.91, b=128.91, c=50.70 α=90°, β=90°, γ=120°
resolution limit (Å)	31 - 2.2	50 – 2.1	50 -1.8
observed reflections	4117	35848	82387
completeness (%)*	98.8 (98.3)	98.2 (94.7)	99.4 (96.5)
redundancy (%)*	6.10 (6.38)	5.63 (5.13)	6.3 (6.0)
R_{sym} (% on I)*	5.8 (13.1)	9.1 (39.4)	6.6 (22.7)
Refinement			
resolution range (Å)	31 - 2.2	50 – 2.1	46 -1.8
R_{work} (%)	16.81	17.53	21.68
R_{free} (%)	22.24	24.00	24.97
ordered water molecules	53	175	257
protein atoms	684	4742	3583
sulphates	1	0	9
methanol	1	0	0
rmsd of bond lengths	0.008	0.002	0.004
rmsd of bond angles	0.944	0.460	0.803
average B-factor	23.7	31.65	31.66
Ramachandran plot			
most favored (%)	95.0	90.0	91.2
additional allowed (%)	5.0	10.0	8.8
generously allowed (%)	0.0	0.0	0.0
disallowed (%)	0.0	0.0	0.0
Crystallization conditions	0.05 M sodium succinate, pH 4.0 2.9 M ammonium sulphate 30% methanol	0.1M sodium citrate, pH 4.1 1.2 M lithium chloride 30 % PEG 6000	0.1 M HEPES, pH 5.5 0.1 M L-cysteine 2.9 M ammonium sulphate

* Values in parenthesis refer to the outermost resolution shell.

Figures

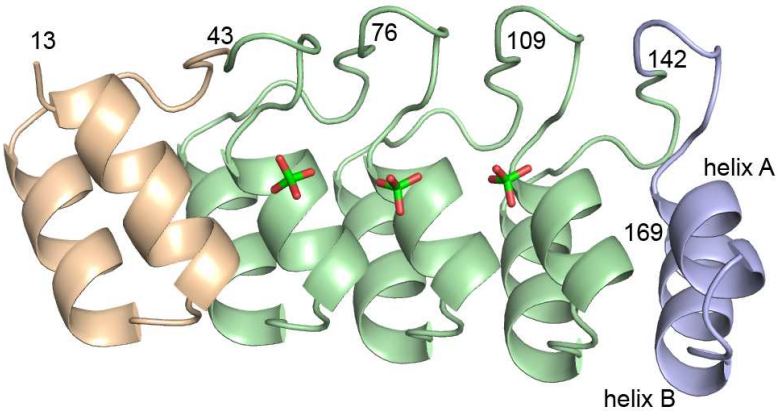


a.

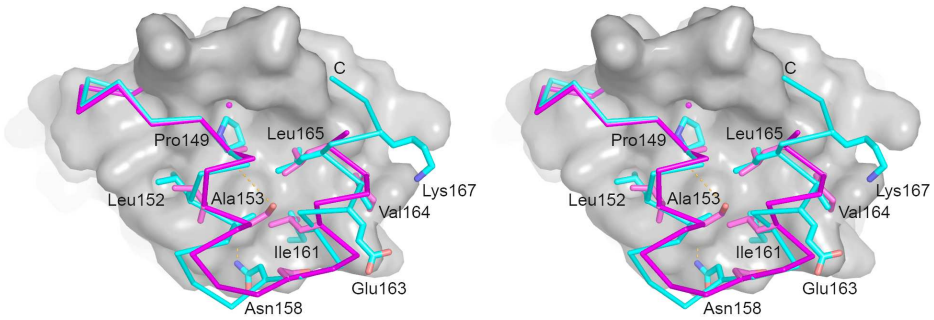


b.

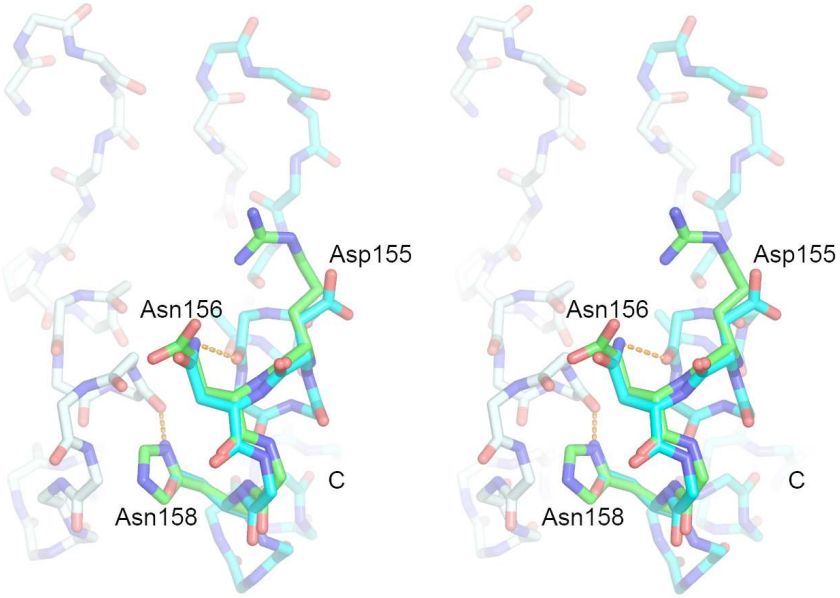
Figure 1.


13
43
76
109
142
helix A
169
helix B

a.


C
Pro149
Leu165
Leu152
Ala153
Lys167
Val164
Ile161
Glu163
Asn158
C
Pro149
Leu165
Leu152
Ala153
Lys167
Val164
Ile161
Glu163
Asn158

b.



c.

Figure 2.

References

[1] Andrade, M.A., Perez-Iratxeta, C. and Ponting, C.P. (2001). Protein repeats: structures, functions, and evolution. J Struct Biol 134, 117-31.
[2] Bork, P. (1993). Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally? Proteins 17, 363-74.
[3] Stumpp, M.T., Forrer, P., Binz, H.K. and Pluckthun, A. (2003). Designing repeat proteins: modular leucine-rich repeat protein libraries based on the mammalian ribonuclease inhibitor family. Journal of Molecular Biology 332, 471-87.
[4] Parmeggiani, F., Pellarin, R., Larsen, A.P., Varadamsetty, G., Stumpp, M.T., Zerbe, O., Caflisch, A. and Pluckthun, A. (2008). Designed armadillo repeat proteins as general peptide-binding scaffolds: consensus design and computational optimization of the hydrophobic core. J Mol Biol 376, 1282-304.
[5] Main, E.R., Xiong, Y., Cocco, M.J., D'Andrea, L. and Regan, L. (2003). Design of stable alpha-helical arrays from an idealized TPR motif. Structure 11, 497-508.
[6] Mosavi, L.K., Minor, D.L., Jr. and Peng, Z.Y. (2002). Consensus-derived structural determinants of the ankyrin repeat motif. Proceedings of the National Academy of Sciences of the United States of America 99, 16029-34.
[7] Binz, H.K., Stumpp, M.T., Forrer, P., Amstutz, P. and Plückthun, A. (2003). Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. Journal of Molecular Biology 332, 489-503.
[8] Kohl, A., Binz, H.K., Forrer, P., Stumpp, M.T., Plückthun, A. and Grütter, M.G. (2003). Designed to be stable: crystal structure of a consensus ankyrin repeat protein. Proceedings of the National Academy of Sciences of the United States of America 100, 1700-5.
[9] Amstutz, P., Binz, H.K., Parizek, P., Stumpp, M.T., Kohl, A., Grütter, M.G., Forrer, P. and Plückthun, A. (2005). Intracellular kinase inhibitors selected from combinatorial libraries of designed ankyrin repeat proteins. Journal of Biological Chemistry 280, 24715-22.
[10] Kohl, A. et al. (2005). Allosteric inhibition of aminoglycoside phosphotransferase by a designed ankyrin repeat protein. Structure 13, 1131-41.
[11] Wetzel, S.K., Settanni, G., Kenig, M., Binz, H.K. and Pluckthun, A. (2008). Folding and unfolding mechanism of highly stable full-consensus ankyrin repeat proteins. J Mol Biol 376, 241-57.
[12] Merz, T., Wetzel, S.K., Firbank, S., Pluckthun, A., Grutter, M.G. and Mittl, P.R. (2008). Stabilizing ionic interactions in a full-consensus ankyrin repeat protein. J Mol Biol 376, 232-40.
[13] Binz, H.K., Kohl, A., Pluckthun, A. and Grutter, M.G. (2006). Crystal structure of a consensus-designed ankyrin repeat protein: implications for stability. Proteins 65, 280-4.
[14] Interlandi, G., Wetzel, S.K., Settanni, G., Pluckthun, A. and Caflisch, A. (2008). Characterization and further stabilization of designed ankyrin repeat proteins by combining molecular dynamics simulations and experiments. J Mol Biol 375, 837-54.

[15] Binz, H.K., Amstutz, P., Kohl, A., Stumpp, M.T., Briand, C., Forrer, P., Grütter, M.G. and Plückthun, A. (2004). High-affinity binders selected from designed ankyrin repeat protein libraries. Nature Biotechnology 22, 575-82.
[16] Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. Journal of Applied Crystallography 26, 795-800.
[17] Adams, P.D. et al. (2002). PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr 58, 1948-54.
[18] McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr 40, 658-674.
[19] Kelly, L., McDonough, M.A., Coleman, M.L., Ratcliffe, P.J. and Schofield, C.J. (2009). Asparagine beta-hydroxylation stabilizes the ankyrin repeat domain fold. Mol Biosyst 5, 52-8.
[20] Emsley, P. and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60, 2126-32.
[21] Davis, I.W. et al. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res 35, W375-83.
[22] Laskowski, R.A., Moss, D.S. and Thornton, J.M. (1993). Main-chain bond lengths and bond angles in protein structures. J Mol Biol 231, 1049-67.
[23] Lawrence, M.C.C., P.M. (1993). Shape complementarity at protein/protein interfaces. J. Mol. Biol. 234, 946-950.
[24] DeLano, W.L. (2002) The PyMOL molecular graphics systemed.^eds). DeLano Scientific, San Carlos, CA, USA.

D) Appendices

Medium for bacterial growth

LB media

10g Bacto-tryptone

5g Bacto-yeast extract

10g NaCl

1 l H_2O

2YT media

16g Bacto-tryptone

10g yeast extract

5g NaCl

1l H_2O

MacKinnon Screens1

	Na-Acetate, #036, 50mM, pH 4.50	Na-Cacodyl., #037, 50mM, pH 5.50	ADA, #038, 50mM, pH 6.50	HEPES, #039, 50mM, pH 7.40	TRIS, #040, 50mM, pH 8.40	Glycine, #041, 50mM, pH 9.40	Na-Acetate, #042, 50mM, pH 4.50	Na-Cacodyl., #043, 50mM, pH 5.50	ADA, #044, 50mM, pH 6.50	HEPES, #045, 50mM, pH 7.40	TRIS, #046, 50mM, pH 8.40	Glycine, #047, 50mM, pH 9.40
	1.0 M NaCl						150 mM NaCl					
#048 PEG 400 15 %	A:1 X: Y: Z:	A:2 X: Y: Z:	A:3 X: Y: Z:	A:4 X: Y: Z:	A:5 X: Y: Z:	A:6 X: Y: Z:	A:7 X: Y: Z:	A:8 X: Y: Z:	A:9 X: Y: Z:	A:10 X: Y: Z:	A:11 X: Y: Z:	A:12 X: Y: Z:
#048 PEG 400 25 %	B:1 X: Y: Z:	B:2 X: Y: Z:	B:3 X: Y: Z:	B:4 X: Y: Z:	B:5 X: Y: Z:	B:6 X: Y: Z:	B:7 X: Y: Z:	B:8 X: Y: Z:	B:9 X: Y: Z:	B:10 X: Y: Z:	B:11 X: Y: Z:	B:12 X: Y: Z:
#048 PEG 400 35 %	C:1 X: Y: Z:	C:2 X: Y: Z:	C:3 X: Y: Z:	C:4 X: Y: Z:	C:5 X: Y: Z:	C:6 X: Y: Z:	C:7 X: Y: Z:	C:8 X: Y: Z:	C:9 X: Y: Z:	C:10 X: Y: Z:	C:11 X: Y: Z:	C:12 X: Y: Z:
#048 PEG 400 45 %	D:1 X: Y: Z:	D:2 X: Y: Z:	D:3 X: Y: Z:	D:4 X: Y: Z:	D:5 X: Y: Z:	D:6 X: Y: Z:	D:7 X: Y: Z:	D:8 X: Y: Z:	D:9 X: Y: Z:	D:10 X: Y: Z:	D:11 X: Y: Z:	D:12 X: Y: Z:
#012 PEG 4000 5 %	E:1 X: Y: Z:	E:2 X: Y: Z:	E:3 X: Y: Z:	E:4 X: Y: Z:	E:5 X: Y: Z:	E:6 X: Y: Z:	E:7 X: Y: Z:	E:8 X: Y: Z:	E:9 X: Y: Z:	E:10 X: Y: Z:	E:11 X: Y: Z:	E:12 X: Y: Z:
#012 PEG 4000 10 %	F:1 X: Y: Z:	F:2 X: Y: Z:	F:3 X: Y: Z:	F:4 X: Y: Z:	F:5 X: Y: Z:	F:6 X: Y: Z:	F:7 X: Y: Z:	F:8 X: Y: Z:	F:9 X: Y: Z:	F:10 X: Y: Z:	F:11 X: Y: Z:	F:12 X: Y: Z:
#012 PEG 4000 20 %	G:1 X: Y: Z:	G:2 X: Y: Z:	G:3 X: Y: Z:	G:4 X: Y: Z:	G:5 X: Y: Z:	G:6 X: Y: Z:	G:7 X: Y: Z:	G:8 X: Y: Z:	G:9 X: Y: Z:	G:10 X: Y: Z:	G:11 X: Y: Z:	G:12 X: Y: Z:
#012 PEG 4000 30 %	H:1 X: Y: Z:	H:2 X: Y: Z:	H:3 X: Y: Z:	H:4 X: Y: Z:	H:5 X: Y: Z:	H:6 X: Y: Z:	H:7 X: Y: Z:	H:8 X: Y: Z:	H:9 X: Y: Z:	H:10 X: Y: Z:	H:11 X: Y: Z:	H:12 X: Y: Z:

MacKinnon Screen 2

	Na-Acetate, #049, 50mM, pH 4.50	Na-Cacodyl., #050, 50mM, pH 5.50	ADA, #051, 50mM, pH 6.5	HEPES, #052, 50mM, pH 7.40	TRIS, #053, 50mM, pH 8.40	Glycine, #054, 50mM, pH 9.40	Na-Acetate, #055, 50mM, pH 4.50	Na-Cacodyl., #056, 50mM, pH 5.50	ADA, #057, 50mM, pH 6.5	HEPES, #058, 50mM, pH 7.40	TRIS, #059, 50mM, pH 8.40	Glycine, #060, 50mM, pH 9.40
	200 mM $(NH_4)_2SO_4$						50 mM Mg-Acetate					
#048 PEG 400 15 %	A:1 X: Y: Z:	A:2 X: Y: Z:	A:3 X: Y: Z:	A:4 X: Y: Z:	A:5 X: Y: Z:	A:6 X: Y: Z:	A:7 X: Y: Z:	A:8 X: Y: Z:	A:9 X: Y: Z:	A:10 X: Y: Z:	A:11 X: Y: Z:	A:12 X: Y: Z:
#048 PEG 400 25 %	B:1 X: Y: Z:	B:2 X: Y: Z:	B:3 X: Y: Z:	B:4 X: Y: Z:	B:5 X: Y: Z:	B:6 X: Y: Z:	B:7 X: Y: Z:	B:8 X: Y: Z:	B:9 X: Y: Z:	B:10 X: Y: Z:	B:11 X: Y: Z:	B:12 X: Y: Z:
#048 PEG 400 35 %	C:1 X: Y: Z:	C:2 X: Y: Z:	C:3 X: Y: Z:	C:4 X: Y: Z:	C:5 X: Y: Z:	C:6 X: Y: Z:	C:7 X: Y: Z:	C:8 X: Y: Z:	C:9 X: Y: Z:	C:10 X: Y: Z:	C:11 X: Y: Z:	C:12 X: Y: Z:
#048 PEG 400 45 %	D:1 X: Y: Z:	D:2 X: Y: Z:	D:3 X: Y: Z:	D:4 X: Y: Z:	D:5 X: Y: Z:	D:6 X: Y: Z:	D:7 X: Y: Z:	D:8 X: Y: Z:	D:9 X: Y: Z:	D:10 X: Y: Z:	D:11 X: Y: Z:	D:12 X: Y: Z:
#012 PEG 4000 5 %	E:1 X: Y: Z:	E:2 X: Y: Z:	E:3 X: Y: Z:	E:4 X: Y: Z:	E:5 X: Y: Z:	E:6 X: Y: Z:	E:7 X: Y: Z:	E:8 X: Y: Z:	E:9 X: Y: Z:	E:10 X: Y: Z:	E:11 X: Y: Z:	E:12 X: Y: Z:
#012 PEG 4000 10 %	F:1 X: Y: Z:	F:2 X: Y: Z:	F:3 X: Y: Z:	F:4 X: Y: Z:	F:5 X: Y: Z:	F:6 X: Y: Z:	F:7 X: Y: Z:	F:8 X: Y: Z:	F:9 X: Y: Z:	F:10 X: Y: Z:	F:11 X: Y: Z:	F:12 X: Y: Z:
#012 PEG 4000 20 %	G:1 X: Y: Z:	G:2 X: Y: Z:	G:3 X: Y: Z:	G:4 X: Y: Z:	G:5 X: Y: Z:	G:6 X: Y: Z:	G:7 X: Y: Z:	G:8 X: Y: Z:	G:9 X: Y: Z:	G:10 X: Y: Z:	G:11 X: Y: Z:	G:12 X: Y: Z:
#012 PEG 4000 30 %	H:1 X: Y: Z:	H:2 X: Y: Z:	H:3 X: Y: Z:	H:4 X: Y: Z:	H:5 X: Y: Z:	H:6 X: Y: Z:	H:7 X: Y: Z:	H:8 X: Y: Z:	H:9 X: Y: Z:	H:10 X: Y: Z:	H:11 X: Y: Z:	H:12 X: Y: Z:

Mac Kinnon Screen 3

	Na-Acetate, #061+#073, 50mM, pH 4.50	Na-Cacodyl., #062+#074, 50mM, pH 5.50	ADA, #063+#075, 50mM, pH 6.50	HEPES, #064+#076, 50mM, pH 7.40	TRIS, #065+#077, 50mM, pH 8.40	Glycine, #066+#078, 50mM, pH 9.40	Na-Acetate, #067+#079, 50mM, pH 4.50	Na-Cacodyl., #068+#080, 50mM, pH 5.50	ADA, #069+#081, 50mM, pH 6.50	HEPES, #070+#082, 50mM, pH 7.40	TRIS, #071+#083, 50mM, pH 8.40	Glycine, #072+#084, 50mM, pH 9.40
	Potassium Phosphate: A1-B6+E1-F6: 200mM; C1-D6+G1-H6 100mM						Lithium Sulfate + Sodium Sulfate: A7-B12+E7-F12: each 100mM; C7-D12+G7-H12 each 50mM					
#048 PEG 400 15 %	A:1 X: Y: Z:	A:2 X: Y: Z:	A:3 X: Y: Z:	A:4 X: Y: Z:	A:5 X: Y: Z:	A:6 X: Y: Z:	A:7 X: Y: Z:	A:8 X: Y: Z:	A:9 X: Y: Z:	A:10 X: Y: Z:	A:11 X: Y: Z:	A:12 X: Y: Z:
#048 PEG 400 25 %	B:1 X: Y: Z:	B:2 X: Y: Z:	B:3 X: Y: Z:	B:4 X: Y: Z:	B:5 X: Y: Z:	B:6 X: Y: Z:	B:7 X: Y: Z:	B:8 X: Y: Z:	B:9 X: Y: Z:	B:10 X: Y: Z:	B:11 X: Y: Z:	B:12 X: Y: Z:
#048 PEG 400 35 %	C:1 X: Y: Z:	C:2 X: Y: Z:	C:3 X: Y: Z:	C:4 X: Y: Z:	C:5 X: Y: Z:	C:6 X: Y: Z:	C:7 X: Y: Z:	C:8 X: Y: Z:	C:9 X: Y: Z:	C:10 X: Y: Z:	C:11 X: Y: Z:	C:12 X: Y: Z:
#048 PEG 400 45 %	D:1 X: Y: Z:	D:2 X: Y: Z:	D:3 X: Y: Z:	D:4 X: Y: Z:	D:5 X: Y: Z:	D:6 X: Y: Z:	D:7 X: Y: Z:	D:8 X: Y: Z:	D:9 X: Y: Z:	D:10 X: Y: Z:	D:11 X: Y: Z:	D:12 X: Y: Z:
#012 PEG 4000 5 %	E:1 X: Y: Z:	E:2 X: Y: Z:	E:3 X: Y: Z:	E:4 X: Y: Z:	E:5 X: Y: Z:	E:6 X: Y: Z:	E:7 X: Y: Z:	E:8 X: Y: Z:	E:9 X: Y: Z:	E:10 X: Y: Z:	E:11 X: Y: Z:	E:12 X: Y: Z:
#012 PEG 4000 10 %	F:1 X: Y: Z:	F:2 X: Y: Z:	F:3 X: Y: Z:	F:4 X: Y: Z:	F:5 X: Y: Z:	F:6 X: Y: Z:	F:7 X: Y: Z:	F:8 X: Y: Z:	F:9 X: Y: Z:	F:10 X: Y: Z:	F:11 X: Y: Z:	F:12 X: Y: Z:
#012 PEG 4000 20 %	G:1 X: Y: Z:	G:2 X: Y: Z:	G:3 X: Y: Z:	G:4 X: Y: Z:	G:5 X: Y: Z:	G:6 X: Y: Z:	G:7 X: Y: Z:	G:8 X: Y: Z:	G:9 X: Y: Z:	G:10 X: Y: Z:	G:11 X: Y: Z:	G:12 X: Y: Z:
#012 PEG 4000 30 %	H:1 X: Y: Z:	H:2 X: Y: Z:	H:3 X: Y: Z:	H:4 X: Y: Z:	H:5 X: Y: Z:	H:6 X: Y: Z:	H:7 X: Y: Z:	H:8 X: Y: Z:	H:9 X: Y: Z:	H:10 X: Y: Z:	H:11 X: Y: Z:	H:12 X: Y: Z:

MacKinnon Screen 4

	Na-Acetate, #085, 50mM, pH 4.50	Na-Cacodyl., #086, 50mM, pH 5.50	ADA, #087, 50mM, pH 6.50	HEPES, #088, 50mM, pH 7.40	TRIS, #089, 50mM, pH 8.40	Glycine, #090, 50mM, pH 9.40	Na-Acetate, #091, 50mM, pH 4.50	Na-Cacodyl., #092, 50mM, pH 5.50	MES, #93, 100mM, pH 6.40	HEPES, #094, 50mM, pH 7.40	TRIS, #095, 50mM, pH 8.40	Glycine, #096, 50mM, pH 9.40
	500 mM KCl						200 mM $CaCl_2$					
#048 PEG 400 15 %	A:1 X: Y: Z:	A:2 X: Y: Z:	A:3 X: Y: Z:	A:4 X: Y: Z:	A:5 X: Y: Z:	A:6 X: Y: Z:	A:7 X: Y: Z:	A:8 X: Y: Z:	A:9 X: Y: Z:	A:10 X: Y: Z:	A:11 X: Y: Z:	A:12 X: Y: Z:
#048 PEG 400 25 %	B:1 X: Y: Z:	B:2 X: Y: Z:	B:3 X: Y: Z:	B:4 X: Y: Z:	B:5 X: Y: Z:	B:6 X: Y: Z:	B:7 X: Y: Z:	B:8 X: Y: Z:	B:9 X: Y: Z:	B:10 X: Y: Z:	B:11 X: Y: Z:	B:12 X: Y: Z:
#048 PEG 400 35 %	C:1 X: Y: Z:	C:2 X: Y: Z:	C:3 X: Y: Z:	C:4 X: Y: Z:	C:5 X: Y: Z:	C:6 X: Y: Z:	C:7 X: Y: Z:	C:8 X: Y: Z:	C:9 X: Y: Z:	C:10 X: Y: Z:	C:11 X: Y: Z:	C:12 X: Y: Z:
#048 PEG 400 45 %	D:1 X: Y: Z:	D:2 X: Y: Z:	D:3 X: Y: Z:	D:4 X: Y: Z:	D:5 X: Y: Z:	D:6 X: Y: Z:	D:7 X: Y: Z:	D:8 X: Y: Z:	D:9 X: Y: Z:	D:10 X: Y: Z:	D:11 X: Y: Z:	D:12 X: Y: Z:
#012 PEG 4000 5 %	E:1 X: Y: Z:	E:2 X: Y: Z:	E:3 X: Y: Z:	E:4 X: Y: Z:	E:5 X: Y: Z:	E:6 X: Y: Z:	E:7 X: Y: Z:	E:8 X: Y: Z:	E:9 X: Y: Z:	E:10 X: Y: Z:	E:11 X: Y: Z:	E:12 X: Y: Z:
#012 PEG 4000 10 %	F:1 X: Y: Z:	F:2 X: Y: Z:	F:3 X: Y: Z:	F:4 X: Y: Z:	F:5 X: Y: Z:	F:6 X: Y: Z:	F:7 X: Y: Z:	F:8 X: Y: Z:	F:9 X: Y: Z:	F:10 X: Y: Z:	F:11 X: Y: Z:	F:12 X: Y: Z:
#012 PEG 4000 20 %	G:1 X: Y: Z:	G:2 X: Y: Z:	G:3 X: Y: Z:	G:4 X: Y: Z:	G:5 X: Y: Z:	G:6 X: Y: Z:	G:7 X: Y: Z:	G:8 X: Y: Z:	G:9 X: Y: Z:	G:10 X: Y: Z:	G:11 X: Y: Z:	G:12 X: Y: Z:
#012 PEG 4000 30 %	H:1 X: Y: Z:	H:2 X: Y: Z:	H:3 X: Y: Z:	H:4 X: Y: Z:	H:5 X: Y: Z:	H:6 X: Y: Z:	H:7 X: Y: Z:	H:8 X: Y: Z:	H:9 X: Y: Z:	H:10 X: Y: Z:	H:11 X: Y: Z:	H:12 X: Y: Z:

MacKinnon Screen 5

	Na-Acetate, #097, 50mM, pH 4.50	Na-Cacodyl., #098, 50mM, pH 5.50	ADA, #099, 50mM, pH 6.50	HEPES, #100, 50mM, pH 7.40	TRIS, #101, 50mM, pH 8.40	Glycine, #102, 50mM, pH 9.40	Na-Acetate, #103, 50mM, pH 4.50	Na-Cacodyl., #104, 50mM, pH 5.50	ADA, #105, 50mM, pH 6.50	HEPES, #106, 50mM, pH 7.40	TRIS, #107, 50mM, pH 8.40	Glycine, #108, 50mM, pH 9.40
	1.0 M Ammonium Formate						50mM Zinc-Acetate			100mM K-Iodide		
#048 PEG 400 15 %	A:1 X: Y: Z:	A:2 X: Y: Z:	A:3 X: Y: Z:	A:4 X: Y: Z:	A:5 X: Y: Z:	A:6 X: Y: Z:	A:7 X: Y: Z:	A:8 X: Y: Z:	A:9 X: Y: Z:	A:10 X: Y: Z:	A:11 X: Y: Z:	A:12 X: Y: Z:
#048 PEG 400 25 %	B:1 X: Y: Z:	B:2 X: Y: Z:	B:3 X: Y: Z:	B:4 X: Y: Z:	B:5 X: Y: Z:	B:6 X: Y: Z:	B:7 X: Y: Z:	B:8 X: Y: Z:	B:9 X: Y: Z:	B:10 X: Y: Z:	B:11 X: Y: Z:	B:12 X: Y: Z:
#048 PEG 400 35 %	C:1 X: Y: Z:	C:2 X: Y: Z:	C:3 X: Y: Z:	C:4 X: Y: Z:	C:5 X: Y: Z:	C:6 X: Y: Z:	C:7 X: Y: Z:	C:8 X: Y: Z:	C:9 X: Y: Z:	C:10 X: Y: Z:	C:11 X: Y: Z:	C:12 X: Y: Z:
#048 PEG 400 45 %	D:1 X: Y: Z:	D:2 X: Y: Z:	D:3 X: Y: Z:	D:4 X: Y: Z:	D:5 X: Y: Z:	D:6 X: Y: Z:	D:7 X: Y: Z:	D:8 X: Y: Z:	D:9 X: Y: Z:	D:10 X: Y: Z:	D:11 X: Y: Z:	D:12 X: Y: Z:
#012 PEG 4000 5 %	E:1 X: Y: Z:	E:2 X: Y: Z:	E:3 X: Y: Z:	E:4 X: Y: Z:	E:5 X: Y: Z:	E:6 X: Y: Z:	E:7 X: Y: Z:	E:8 X: Y: Z:	E:9 X: Y: Z:	E:10 X: Y: Z:	E:11 X: Y: Z:	E:12 X: Y: Z:
#012 PEG 4000 10 %	F:1 X: Y: Z:	F:2 X: Y: Z:	F:3 X: Y: Z:	F:4 X: Y: Z:	F:5 X: Y: Z:	F:6 X: Y: Z:	F:7 X: Y: Z:	F:8 X: Y: Z:	F:9 X: Y: Z:	F:10 X: Y: Z:	F:11 X: Y: Z:	F:12 X: Y: Z:
#012 PEG 4000 20 %	G:1 X: Y: Z:	G:2 X: Y: Z:	G:3 X: Y: Z:	G:4 X: Y: Z:	G:5 X: Y: Z:	G:6 X: Y: Z:	G:7 X: Y: Z:	G:8 X: Y: Z:	G:9 X: Y: Z:	G:10 X: Y: Z:	G:11 X: Y: Z:	G:12 X: Y: Z:
#012 PEG 4000 30 %	H:1 X: Y: Z:	H:2 X: Y: Z:	H:3 X: Y: Z:	H:4 X: Y: Z:	H:5 X: Y: Z:	H:6 X: Y: Z:	H:7 X: Y: Z:	H:8 X: Y: Z:	H:9 X: Y: Z:	H:10 X: Y: Z:	H:11 X: Y: Z:	H:12 X: Y: Z:

MacKinnon Screen 6

	Na-Acetate, #115, 50mM, pH 4.50	Na-Cacodyl., #116, 50mM, pH 5.50	ADA, #117, 50mM, pH 6.50	HEPES, #118, 50mM, pH 7.40	TRIS, #119, 50mM, pH 8.40	Glycine, #120, 50mM, pH 9.40	MES, #109, 100mM, pH 6.40	Na-Cacodyl., #110, 50mM, pH 5.50	ADA, #111, 50mM, pH 6.50	HEPES, #112, 50mM, pH 7.40	TRIS, #113, 50mM, pH 8.40	Glycine, #114, 50mM, pH 9.40
	No Addition							100mM Potassium-Sodium-Tartrate				
#048 PEG 400 15 %	A:1 X: Y: Z:	A:2 X: Y: Z:	A:3 X: Y: Z:	A:4 X: Y: Z:	A:5 X: Y: Z:	A:6 X: Y: Z:	A:7 X: Y: Z:	A:8 X: Y: Z:	A:9 X: Y: Z:	A:10 X: Y: Z:	A:11 X: Y: Z:	A:12 X: Y: Z:
#048 PEG 400 25 %	B:1 X: Y: Z:	B:2 X: Y: Z:	B:3 X: Y: Z:	B:4 X: Y: Z:	B:5 X: Y: Z:	B:6 X: Y: Z:	B:7 X: Y: Z:	B:8 X: Y: Z:	B:9 X: Y: Z:	B:10 X: Y: Z:	B:11 X: Y: Z:	B:12 X: Y: Z:
#048 PEG 400 35 %	C:1 X: Y: Z:	C:2 X: Y: Z:	C:3 X: Y: Z:	C:4 X: Y: Z:	C:5 X: Y: Z:	C:6 X: Y: Z:	C:7 X: Y: Z:	C:8 X: Y: Z:	C:9 X: Y: Z:	C:10 X: Y: Z:	C:11 X: Y: Z:	C:12 X: Y: Z:
#048 PEG 400 45 %	D:1 X: Y: Z:	D:2 X: Y: Z:	D:3 X: Y: Z:	D:4 X: Y: Z:	D:5 X: Y: Z:	D:6 X: Y: Z:	D:7 X: Y: Z:	D:8 X: Y: Z:	D:9 X: Y: Z:	D:10 X: Y: Z:	D:11 X: Y: Z:	D:12 X: Y: Z:
#012 PEG 4000 5 %	E:1 X: Y: Z:	E:2 X: Y: Z:	E:3 X: Y: Z:	E:4 X: Y: Z:	E:5 X: Y: Z:	E:6 X: Y: Z:	E:7 X: Y: Z:	E:8 X: Y: Z:	E:9 X: Y: Z:	E:10 X: Y: Z:	E:11 X: Y: Z:	E:12 X: Y: Z:
#012 PEG 4000 10 %	F:1 X: Y: Z:	F:2 X: Y: Z:	F:3 X: Y: Z:	F:4 X: Y: Z:	F:5 X: Y: Z:	F:6 X: Y: Z:	F:7 X: Y: Z:	F:8 X: Y: Z:	F:9 X: Y: Z:	F:10 X: Y: Z:	F:11 X: Y: Z:	F:12 X: Y: Z:
#012 PEG 4000 20 %	G:1 X: Y: Z:	G:2 X: Y: Z:	G:3 X: Y: Z:	G:4 X: Y: Z:	G:5 X: Y: Z:	G:6 X: Y: Z:	G:7 X: Y: Z:	G:8 X: Y: Z:	G:9 X: Y: Z:	G:10 X: Y: Z:	G:11 X: Y: Z:	G:12 X: Y: Z:
#012 PEG 4000 30 %	H:1 X: Y: Z:	H:2 X: Y: Z:	H:3 X: Y: Z:	H:4 X: Y: Z:	H:5 X: Y: Z:	H:6 X: Y: Z:	H:7 X: Y: Z:	H:8 X: Y: Z:	H:9 X: Y: Z:	H:10 X: Y: Z:	H:11 X: Y: Z:	H:12 X: Y: Z:

Additive Screen

Additive	10 x [c]
Cadmium chloride	100mM
Calcium chloride dehydrate	100mM
Cobalt chloride hexahydrate	100mM
Magnesium chloride hexahydrate	100mM
Manganese- (II)-chloride tetrahydrate	100mM
Barium chloride dehydrate	100mM
Cupric- (II)-chloride, dihydrate	100mM
Zinc chloride	100mM
Samarium- (III)-chloride	100mM
Holmium- (III)-chloride	100mM
Lithium chloride anhydrous	1.0M
Sodium fluoride	0.5M
Thallium- (I)-chloride	10mM
Ammonium sulfate	1M
Cesium chloride	1M
Potassium chloride	1M
Sodium chloride	2M
Nickel- (II)-sulfate hexahydrate	100mM
Ferric- (III)-chloride	100mM
Sodium molybdate dihydrat	100mM
Potassium bromide	100mM
Malonic Acid	1M
Lead- (II)-acetate trihydrate	100mM
Sodium iodide	1M
Xylitol	30% (w/v)
D (+)-Glucose anhydrous	30% (w/v)
D (+)-Sucrose	30% (w/v)
Ascorbic Acid	10% (w/v)
D-Sorbitol	30% (w/v)
D-Mannitol	15% (w/v)
D (-)-Fructose	30% (w/v)
Boric Acid	500mM
Guanidine HCl	1M
Urea	100mM
DMSO	30% (v/v)
1,1'-Ferrocenedicarboxylic acid	125mM
EDTA disodium salt dihydrate	100mM
Nicotinamid	100mM
Oxalic Acid	100mM
Rubidium Chloride	100mM
6-Aminohexanoic acid	30% (w/v)
Trehalose Dihydrate	10% (w/v)
1,6-Diaminohexane	28% (w/v)
Taurin	100mM
Betaine monohydrate	100mM
Glycine	1M

Additive	10 x [c]
LDAO	1.4mM
Thiamine HCl	100mM
Ethylene glycol	30% (v/v)
Glycerol 87%	30% (w/v)
1,6-Hexanediol	30% (w/v)
2-Methyl-2,4-pentanediol	30% (v/v)
PEG 400	50% (v/v)
Jeffamine M600 pH~7	50% (v/v)
PEG 6000	15% (w/v)
PEG 20,000	15% (w/v)
PEG 2000 MME	20% (w/v)
Succinic Acid	1M
1-Propanol (n-Propanol)	40% (v/v)
Terephthalic acid	500mM
Sodium benzoate	500mM
Phthalic acid	500mM
Tert. Butanol	40% (v/v)
Dioxane	30% (v/v)
p-Toluic acid	500mM
2-Propanol (i-Propanol)	30% (v/v)
Methanol	30% (v/v)
Phenoxyacetic acid	500mM
1-Naphthoic acid	500mM
Acetonitrile	40% (v/v)
1-Butanol (n-Butanol)	7% (v/v)
PEG 200	40% (v/v)
Spermine	100mM
1.4-Phenylendiamin	500mM
Aniline	250mM
Sodium benzenesulfonate	500mM
DTT (1.4-Dithio-DL-Threitol)	100mM
L-Cysteine	~100mM
L-Glutamic acid	100mM
L-Arginine monohydrochloride	100mM
L-Lysine	100mM
DL-Serine	100mM
L-Glutathione reduced form	100mM
L-Asparagine Monohydrate	100mM
L-Threonine	100mM
L-Valine	100mM
Benzamidine HCl	20% (w/v)
Triton X-100	8.6mM
Congo Red	125mM
Deoxycholic acid sodium salt	100mM
n-Dodecyl-beta-d-maltopyranoside	1.7mM
IPTG	10mM

GS002

	$(NH_4)_2SO_4$	$(NH_4)_2SO_4$	$(NH_4)_2SO_4$	$(NH_4)_2SO_4$	$(NH_4)_2SO_4$	$(NH_4)_2SO_4$	$(NH_4)_2SO_4$	$(NH_4)_2SO_4$	$(NH_4)_2SO_4$	$(NH_4)_2SO_4$	$(NH_4)_2SO_4$	$(NH_4)_2SO_4$
	15 %	20 %	25 %	30 %	35 %	40 %	45 %	50%	55 %	60 %	70 %	80 %
#003 Malonic Acid, 0.1M, pH 3.0	A:1 X: Y: Z:	A:2 X: Y: Z:	A:3 X: Y: Z:	A:4 X: Y: Z:	A:5 X: Y: Z:	A:6 X: Y: Z:	A:7 X: Y: Z:	A:8 X: Y: Z:	A:9 X: Y: Z:	A:10 X: Y: Z:	A:11 X: Y: Z:	A:12 X: Y: Z:
#004 Succinic Acid 0.05M, pH 4.0	B:1 X: Y: Z:	B:2 X: Y: Z:	B:3 X: Y: Z:	B:4 X: Y: Z:	B:5 X: Y: Z:	B:6 X: Y: Z:	B:7 X: Y: Z:	B:8 X: Y: Z:	B:9 X: Y: Z:	B:10 X: Y: Z:	B:11 X: Y: Z:	B:12 X: Y: Z:
#005 Citric Acid 0.1M, pH 5.0	C:1 X: Y: Z:	C:2 X: Y: Z:	C:3 X: Y: Z:	C:4 X: Y: Z:	C:5 X: Y: Z:	C:6 X: Y: Z:	C:7 X: Y: Z:	C:8 X: Y: Z:	C:9 X: Y: Z:	C:10 X: Y: Z:	C:11 X: Y: Z:	C:12 X: Y: Z:
#006 MES 0.1M, pH 6.0	D:1 X: Y: Z:	D:2 X: Y: Z:	D:3 X: Y: Z:	D:4 X: Y: Z:	D:5 X: Y: Z:	D:6 X: Y: Z:	D:7 X: Y: Z:	D:8 X: Y: Z:	D:9 X: Y: Z:	D:10 X: Y: Z:	D:11 X: Y: Z:	D:12 X: Y: Z:
#007 MOPS 0.1M pH 7.0	E:1 X: Y: Z:	E:2 X: Y: Z:	E:3 X: Y: Z:	E:4 X: Y: Z:	E:5 X: Y: Z:	E:6 X: Y: Z:	E:7 X: Y: Z:	E:8 X: Y: Z:	E:9 X: Y: Z:	E:10 X: Y: Z:	E:11 X: Y: Z:	E:12 X: Y: Z:
#008 HEPES 0.1M pH 8.0	F:1 X: Y: Z:	F:2 X: Y: Z:	F:3 X: Y: Z:	F:4 X: Y: Z:	F:5 X: Y: Z:	F:6 X: Y: Z:	F:7 X: Y: Z:	F:8 X: Y: Z:	F:9 X: Y: Z:	F:10 X: Y: Z:	F:11 X: Y: Z:	F:12 X: Y: Z:
#009 TAPS 0.02M pH 9.0	G:1 X: Y: Z:	G:2 X: Y: Z:	G:3 X: Y: Z:	G:4 X: Y: Z:	G:5 X: Y: Z:	G:6 X: Y: Z:	G:7 X: Y: Z:	G:8 X: Y: Z:	G:9 X: Y: Z:	G:10 X: Y: Z:	G:11 X: Y: Z:	G:12 X: Y: Z:
#010 CAPS 0.02M pH 10.0	H:1 X: Y: Z:	H:2 X: Y: Z:	H:3 X: Y: Z:	H:4 X: Y: Z:	H:5 X: Y: Z:	H:6 X: Y: Z:	H:7 X: Y: Z:	H:8 X: Y: Z:	H:9 X: Y: Z:	H:10 X: Y: Z:	H:11 X: Y: Z:	H:12 X: Y: Z: